

虚拟网络行为对互联网整体特性的影响^{*}

王 磊 周淑华 袁 坚 任 勇 山秀明

(清华大学电子工程系, 北京 100084)

(2006 年 4 月 30 日收到; 2006 年 6 月 4 日收到修改稿)

虚拟网络是一种依赖 Internet 基础设施所提供的传输能力, 但又具有独立拓扑结构和信息传递规则的应用层网络行为逻辑网络. 提出了耦合虚拟网络行为与物理节点的抽象模型, 对一类典型的虚拟网络逻辑拓扑给互联网整体特性带来的影响进行了分析. 研究表明在虚拟网络作用下, 节点数据包排队长度存在相变特性, 但相变临界点比对规则网络发生了明显左移, 网络性能相对恶化. 当数据包注入速率小于相变临界速率时, 节点数据包排队长度不相关或短程相关; 在接近临界速率处, 节点数据包排队长度长程相关, 幂指数 H 增大, 网络获得更强的长程相关性. 同时, 在注入速率大于或等于临界速率时, 虚拟网络行为使网络呈现出一致的长程相关特性.

关键词: 虚拟网络, 元胞自动机, 相变, 互联网

PACC: 0540J, 8980J

1. 引 言

旨在揭示互联网 (Internet) 日益加剧的复杂性的网络整体动力学理论与方法的探索取得了长足的进展. Internet 演进中的普适规律和不变特征正被逐步揭示, 如业务量的自相似特性、TTL (数据包生存时间) 分布的 $1/f$ 特性、网络中节点的长程相关特性及处理器利用率的幂律分布等^[1-3], 同时, 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.

从时间演化角度, 引入相变的概念以解释数据包 TTL 和业务量密度的波动, 多数研究表明网络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一致的幂律特性^[4-8]. 进而, 文献[9]通过对节点排队长度均方涨落函数的统计, 研究了节点在时间上的长程相关特性, 并发现只有在突破相变临界状态后, 节点才开始呈现长程相关特性, 即大于 0.5 的幂指数特征. 从空间演化角度, 我们已有的工作通过对数据包排队长度在空间上的互相关函数的分析, 指出节点排队长度的互相关特性中也存在明显的相变现象及 $1/f$ 特性^[10, 11]. 这些基于统计物理学的研究工作初步探索了分组交换网络流量的临界现象, 大大增强了人们对 Internet “突发性” 流量成因的理解.

然而, 这一类解释在网络模型构建中都默许了网络节点独立性、同一性的假设, 而忽略了业务源时空特征的差异, 忽略了应用层网络行为 (用户行为) 的差异及其对网络流量时空分布的影响. Barabási 等人的研究发现万维网节点的度分布具有幂律特性^[12]. 这种统计特征表明在万维网的逻辑拓扑中, 多数用户热衷于访问为数不多的几个热门站点, 映射到 Internet 物理网络, 用户的一致性偏好行为使得网络流量在空间分布极不平衡. 事实上, 在物理网络 (Internet) 之上, 用户行为与 Internet 涉及的应用对象在逻辑上构成了一个上层的“虚拟网络 (virtual networks)”. 虚拟网络改变了 Internet “端到端” 的意义, 使得面向网络的端到端传递过程和面向用户的端到端传递过程相分离, 从而影响到作为物理基础设施的 Internet 承载最终用户通信的能力, 也很可能影响到 Internet 的整体行为特征. 虚拟网络拓扑结构的形成不受物理设施部署的约束, 这种灵活性可能加剧这种影响的程度. 文献[13]研究了用户关系网络与电话网络的互作用, 指出了用户关系网络对电话网络负载性能的显著影响. 与此类似的研究在 Internet 应用场景中并没有给予足够的重视, 虚拟网络行为对互联网整体特性的影响也尚未见诸文献报道.

为了理解虚拟网络和物理网络相互作用对 Internet 整体动力学行为的影响, 本文以二维网络元

胞自动机模型为基础, 借鉴统计物理学在复杂网络结构演化方面的研究成果, 提出一种耦合虚拟网络行为与物理网络的双层模型. 对一典型的虚拟网络——万维网 (World-Wide Web) 的逻辑拓扑对 Internet 整体特性的影响进行了分析. 研究表明在虚拟网络作用下, 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, 网络性能进一步恶化. 利用节点排队长度累积量的均方涨落函数研究了虚拟网络行为对网络长程相关性的影响. 发现当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, 并且发现了幂指数 H 明显增大, 网络的长程相关性得到了增强. 同时, 在注入速率大于或等于临界速率时, 虚拟网络行为使网络呈现出一致的长程相关特性.

2. 二维规则网络抽象模型

对于二维网络元胞自动机抽象模型的研究近期主要集中于对模型边界条件是否满足周期性、网络信源信宿节点与中继节点的划分等问题的讨论与比较^[5,6,14,15]. Fuks 等人^[6]的模型中, 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. 其模型如图 1 所示. 在对 Internet 进行整体建模时, 将其视作为周期边界条件, 便于仿真处理, 而且在理论上也是合理的.

我们已有的工作定义了二维网络模型中数据包的传输规则^[9]. 设 L 为系统的规模 (即每行/列的节点数), 则网络中共有 $N=L^2$ 个节点. 节点的位置用矢量 r 表示为

$$r = i c_x + j c_y, \quad (1)$$

其中 c_x 和 c_y 分别为笛卡儿单位矢量, $i, j=1, \dots, L$. 每个节点以直接相连的四个节点为邻居, 其邻域表示为

$$C(r) = \{r - c_x, r + c_x, r - c_y, r + c_y\}. \quad (2)$$

每一时步的数据包产生、转发和消亡规则如下:

1) 每一时步, 每个节点以一定概率 λ 产生数据包, 其目的节点从其他节点中随机选取, 且新产生的数据包置放在节点数据包队列的最后面.

2) 每一时步, 每个节点从其数据包队列的最前面取出一个数据包, 按照路由策略从该节点的邻域节点中选取一个节点并转发该数据包. 转发过去的数据包排在该邻节点数据包队列的最后面. 如果有

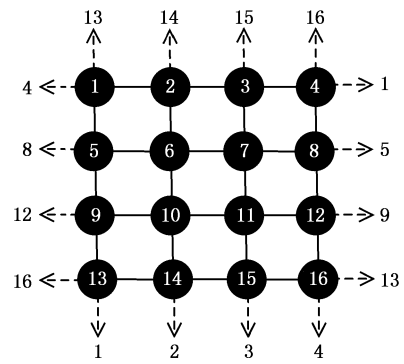


图 1 二维网络模型

几个数据包同时到达同一个节点, 则它们随机排序.

3) 每一时步, 所有数据包的传输时间都增加 1.

4) 如果数据包到达其目的节点, 则该数据包就此消亡.

其中概率 λ 也称作数据包的产生概率或注入概率, 数据包是网络的负载. 各节点都有一个缓存区, 将来不及转发的数据包暂存, 每一时步各节点将排在缓存区最前面的一个数据包传送到 $C(r)$ 中之一. 这里假设节点缓存区足够大.

图 2 所示模型在周期边界条件下, 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^[9]:

$$d_p(r_1, r_2) = L - \left| |i_2 - i_1| + \frac{L}{2} \right| - \left| |j_2 - j_1| + \frac{L}{2} \right|, \quad (3)$$

它和通过图论中的最短路径计算出来的距离结果一致, 但计算更为简单.

数据包路由采用确定性策略, 即如果 $C(r)$ 中离目的节点距离最近的只有一个, 则选择该节点; 如果 $C(r)$ 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. 这实际上是一个全局的最短路径路由策略.

3. 虚拟网络与物理网络的双层耦合模型

虚拟网络是一种依赖 Internet 基础设施所提供的传输能力, 但又具有独立拓扑结构和信息传递规则的应用层网络行为逻辑网络. 一方面, Internet 应用所涉及的对象可能在逻辑上构成网络关系, 如日益庞大的万维网, 这种虚拟网络是在应用发展过程中自然形成的; 另一方面, 人为构造以隧道 (tunnel)

作为基础的虚拟网络, 以实现对 Internet 现有功能的扩展, 如实现分布式对等模式 (peer-to-peer) 的信息检索和跨越地域的企业级互连等.

图 2 以万维网为例说明了虚拟网络的逻辑拓扑与物理网络真实拓扑之间的作用方式. 节点 i_v 对节点 j_v 的访问, 即用户 i_v 点击链接访问站点 j_v 的行为, 在逻辑拓扑中“一跳”即可实现, 而在对应的真实物理网络中, i_p 与 j_p 之间却需要经过 3 次路由转发. 不同于规则网络的抽象模型, 数据包目的节点随机选择的方式被改变, 在统计意义上依照虚拟网络拓扑的指导集中于某些局部区域, 或者说部分负载集中于某些特定的节点.

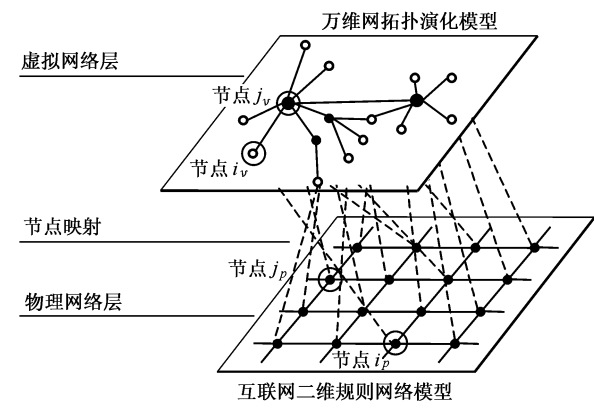


图 2 虚拟网络与物理网络相互作用示意图

为了研究虚拟网络对物理网络的影响, 本文提出一种抽象的双层耦合模型. 在对虚拟网络与物理网络分别进行抽象建模的同时, 耦合其互作用关系. 模型底层为二维元胞自动机模型, 如上文所述, 用来仿真 Internet 物理网络. 顶层为具有幂律特性的万维网拓扑模型, 用来仿真应用层网络行为, 即虚拟网络拓扑分布. 这里使用 Barabási 提出的 BA 模型来模拟万维网拓扑分布, 该模型基于平均场理论, 是网络拓扑尺度无关现象研究的一个里程碑^[12]. 令 $G(t)$ 表示网络随时间的演化结果, 每一时步网络的演化规则如下:

- 1) 初始状态下, 网络中有 m_0 个节点, 相互孤立.
- 2) 网络演化的每一时步, 在原网络 $G(t-1)$ 的基础上加入一个具有 m ($m \leq m_0$) 条边的节点, 这个节点会被随机地分配一个在网络中的位置, 形成新的网络 $G(t)$. m 条边与网络中已经存在的 m 个不同的节点连接.

3) 偏好依附规则: 新加入边的另一个端点按照正比于顶点度数 k 的分布

$$\Pi(k_i) = \frac{k_i}{\sum k_j} \quad (4)$$

随机选取. t 时步后, 形成一个规模为 $N = m_0 + t$ 的网络. 各节点的度满足幂律分布 $p(k) \sim 2m^2 k^{-3}$.

考虑到元胞自动机网络具有周期性边界条件, 采用随机的方式将虚拟网络中的每个节点一一映射到物理网络中去, 如图 2 所示. 需要指出的是, 在实际网络中万维网物理节点大多分布于骨干网的边缘地带, 且呈不规则分布, 仿真着重于通过对模型的抽象探索虚拟网络对物理网络的影响, 故在此处作了简化假设. 在耦合网络节点的基础上, 通过如下的规则耦合网络行为:

1) 每一时步, 物理网络中的每个节点 i_p 以一定概率 λ 产生数据包. 与此同时, 每个节点在虚拟网络中的映射 i_v 在其邻域按照偏好依附规则寻找目的节点 j_v , j_v 在物理网络中的映射 j_p 即为物理网络中数据包的发送地址.

2) 后面步骤与第 2 节中所述相同.

事实上, 上述网络行为耦合机理的本质是通过虚拟网络的作用重新指导分配物理网络中的数据包目的地址, 使得物理网络中数据包的传输过程能够包含应用层的网络行为.

4. 仿真结果及其分析

4.1. 虚拟网络作用下数据包排队长度的整体特性

仿真对规则网络模型与耦合虚拟网络行为的新模型进行了数值模拟与比较. 基于验证仿真结果的一致性, 仿真分别在 25, 100 两个网络规模下进行. 结果如图 3 所示. 其中横轴表示数据包注入速率, 纵轴表示数据包平均队列长度. 其数据为系统经过 500, 1000 时步过渡后得到. 图中两条垂直实线分别表示两种网络模型下的系统临界注入速率.

仿真结果表明, 两种网络模型中节点的平均队列长度均存在相变特性, 虚拟网络行为的作用使相变临界注入速率明显减小, 即使网络相变点发生左移 (图 3(c) 为相变点左移现象的放大效果). 不同网络规模下的仿真均验证了此结论.

对此的解释是, 由于大多的节点遵照虚拟网络的偏好依附策略选择目的节点, 导致在某些“富节

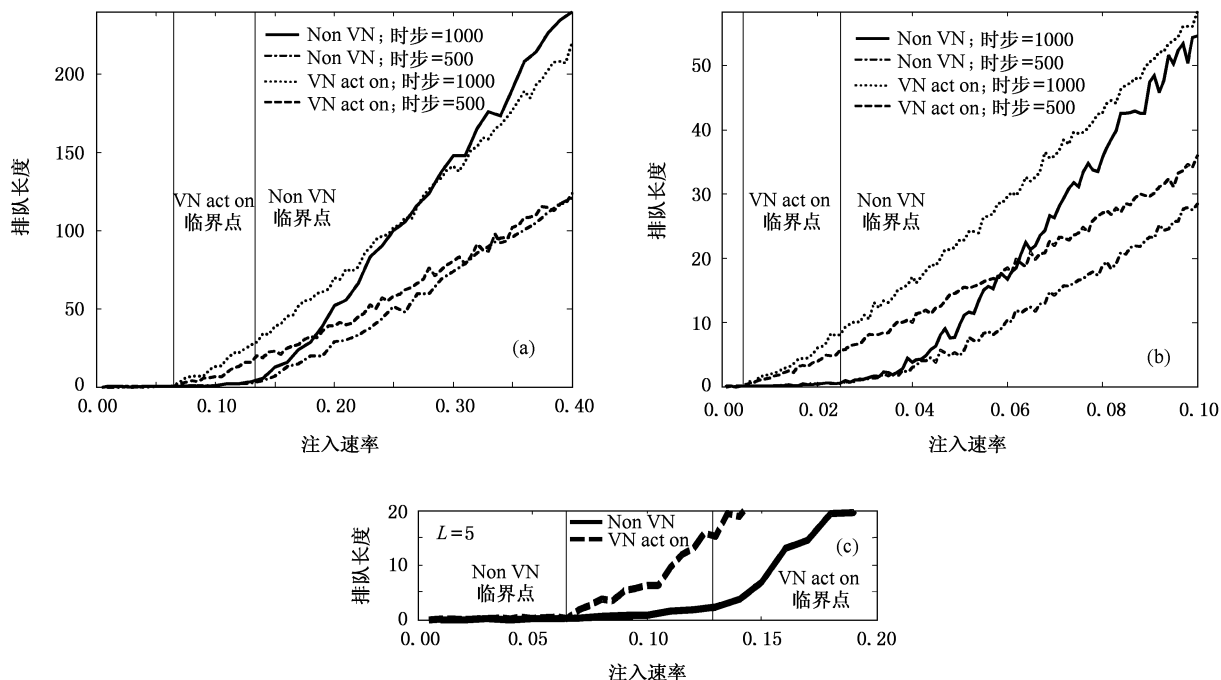


图3 (a)网络规模 25 和(b)网络规模 100 规则网络模型(Non VN)与虚拟网络行为作用后(VN act on)排队长度随数据包注入速率变化对比图(在周期边界条件下,数据为系统经过 500、1000 时步的过渡后得到), (c)网络规模 25 相变点附近曲线放大图

点”周围的局部范围内产生拥塞.而且,由于传输规则(协议),使得网络节点在局部范围发生非线性的相互作用,而这种局部的非线性作用随数据包的传输过程逐渐扩散和影响到整个网络,导致了网络整体性能的相对下降.

若以数据包路由分布来衡量网络整体流量的均衡性,当目的节点随机选取时,网络为均衡网络.虚拟网络行为的引入打破了这种均衡性,随着均衡性的下降,数据包处理能力与网络有效使用率都逐渐减小.已有研究从网络整体特征入手研究网络安全问题^[19],在各角度上,对某些特定节点的恶意攻击则视为一种虚拟网络行为,该行为通过加剧受攻击节点的负载,破坏网络的整体平衡性,导致网络全局性能的下降.

4.2 虚拟网络作用下互联网长程相关特性的变化与分析

文献[9]利用节点排队长度累积量的均方涨落函数,分析了计算机网络中数据包排队长度的长程相关特性.在此工作基础上,本文考察了引入虚拟网络作用后物理网络长程相关特性的相对变化,并探索新机理作用下互联网长程相关特性所呈现出的新统计特征.

令 $q(r, t)$ 表示 t 时刻节点 r 的数据包排队长

度.考虑到虚拟网络与物理网络中节点的随机一一映射,略去节点 r ,则 t 时刻数据包排队长度表示为 $q(t)$.设节点数据包排队长度在时间段 $[1, k]$ 上的累积量为

$$Q(k) = \sum_{t=1}^k q(t), \quad (5)$$

k 为时间间隔.定义其涨落函数为

$$\Delta Q(\tau) \triangleq Q(t_0 + \tau) - Q(t_0), \quad (6)$$

其中 $t_0 = 1, 2, \dots$.定义 $\Delta Q(\tau)$ 涨落的均方根为

$$F(\tau) \triangleq [\overline{\Delta Q(\tau)^2} - \overline{\Delta Q(\tau)}^2]^{0.5}, \quad (7)$$

其中 $\overline{\Delta Q(\tau)}$ 表示 $\Delta Q(\tau)$ 对所有可能的 t_0 平均值. $F(\tau)$ 是一个累积量的均方涨落函数.对一个稳态时间序列而言,涨落函数 $F(\tau)$ 和 τ 应满足如下的幂律关系:

$$F(\tau) \propto \tau^H \quad (8)$$

其中, H 为 Hurst 指数.当 $H = 0.5$ 时,表示 $q(t)$ 不相关或短程相关;当 $H \neq 0.5$ 时,表示时间序列长程相关(long-range correlation)^[17].为验证耦合虚拟网络行为后节点数据包排队长度在时间上的相关特性,本文根据实际的计算结果和理论上的结论进行比较,进而判断实际的相关特性.针对不同的网络规模,仿真结果如图 4 所示.

耦合虚拟网络行为后,当网络中数据包注入速

率为 0.5 倍临界注入速率时, 节点平均队列长度趋近于零, 该时间序列几乎不相关或弱相关. 与之对应, $F(\tau)-\tau$ 在双对数坐标系中曲线斜率 $K_{0.5} \approx 0.5$ (比照曲线 $F(\tau)=\tau^{0.5}$). 当数据包注入速率为 0.9 倍临界注入速率时, 数据包排队长度短程相关, 也有 $K_{0.9} \approx 0.5$. 当数据包注入速率达到临界速率时, 平均队列长度开始显著增大, 且开始呈现长程相关特性, $K_c \approx H(0.5 < H < 1$, 如图 4 中虚线所示). 上述分析表明, 在虚拟网络行为的作用下, 数据包注入速率由零至临界速率增大阶段, 互联网长程相关特性与在规则网络模型下的分析结论一致. 即当注入速率小于临界速率时, 节点数据包排队长度不相关或短程相关; 在接近临界速率处, 节点数据包排队长度开始呈现长程相关特性.

值得注意的是, 相对规则网络模型临界注入速

率的 $F(\tau)-\tau$ 曲线, 耦合虚拟网络行为后 H 值有了提升(在图 4 中体现为直线斜率的增大). 这表明, 虚拟网络的作用使互联网业务流量在整体上呈现出了更强的长程相关性. 不同网络规模下的仿真均验证了此结论. 事实上, 互联网中的每个节点既是信源也是信宿, 在数据包从端到端的传递过程中, 节点间的局部相互作用也逐渐扩散到全局. 虚拟网络行为使得数据包传输在时间与空间上具有更为明显的一致性与整体关联性. 随着系统负载的增加, 虚拟网络行为的整体影响逐渐明显, 其对网络节点间相关性的整体作用也逐渐涌现. 因此, 规则网络与耦合虚拟网络后呈现的长程相关特性, 均可以解释为节点间相互作用的结果; 而虚拟网络行为所引发的更强的长程相关性, 则是节点间整体涌现特性的表征, 是一种对互联网更为真实的刻画.

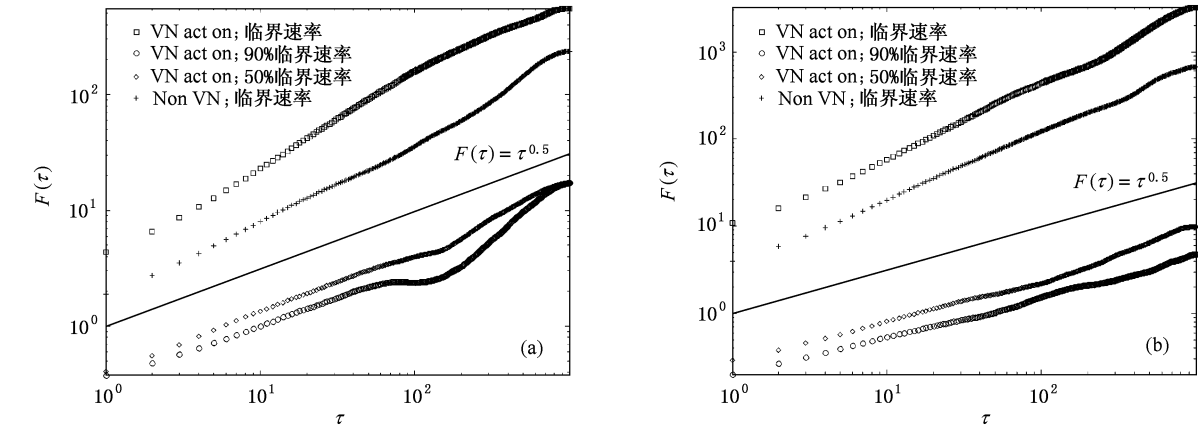


图 4 (a) 网络规模 25 和 (b) 网络规模 100 规则网络模型与虚拟网络行为作用下排队长度在时间上的相关特性图(虚拟网络作用下, 数据包注入速率分别为 1.0(□), 0.9(◇), 0.5(○)倍临界注入速率. 无虚拟网络作用, 为 1.0(+)倍临界注入速率. 图中实线斜率为 0.5)

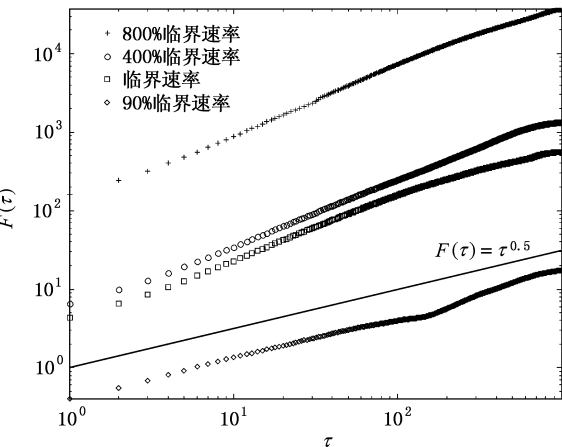


图 5 虚拟网络行为作用下排队长度在时间上的一致相关特性图(数据包注入速率分别为 8(+), 4(◇), 1(□), 0.9(○)倍临界注入速率)

此外, 在相变临界点的右侧空间考察网络的长程相关性时发现, 在数据包注入速率超过临界速率后, 网络性能急剧恶化, 拥塞逐渐加剧并使节点排队长度趋于无穷. 仿真中通过控制系统时步, 得到了 4 倍、8 倍临界注入速率下的 $F(\tau)-\tau$ 曲线. 通过图 5 比对发现, 在虚拟网络行为作用下, 当注入速率到达和超过临界速率, $F(\tau)-\tau$ 曲线具有几乎相同的斜率(H 值), 即网络呈现出一致的长程相关特性. 相对于规则网络模型, 虚拟网络行为使网络节点间数据包的产生消亡具有了更强的一致性. 在相变临界点左侧, 节点间的相互作用逐渐加剧; 直至越过临界点, 虚拟网络拓扑特性开始起主导作用, 使网络呈现出不同于规则网络模型的一致的长程相关特性. 这也是虚拟网络行为整体涌现的一个表征.

5. 结论与讨论

为了理解虚拟网络和物理网络相互作用对 Internet 整体动力学行为的影响, 本文提出了一种耦合虚拟网络行为与物理网络的双层模型, 探索了虚拟网络使网络整体特性产生的变化与新现象: 1) 虚拟网络作用下, 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. 2) 当注入速率小于临界速率时, 与规则网络下具有一致的结论. 3) 在接近临界速率处, 节点数据包排队长度长程相关, 幂指数 H 相对变大, 网络的长程相关性得到了增强. 4) 在注入速率大于或等于临界速率时, 虚拟网络行为使网络呈现出一致的长程相关特性.

上述特征的出现源于虚拟网络行为对物理网络整体流量分布的逻辑驱动, 应用层网络行为的偏好性使负载平衡性降低, 加剧了网络局部恶化速度, 从而影响到整个系统. 网络长程相关性的增强, 进一步印证了虚拟网络行为与互联网节点间的整体涌现特性. 虚拟网络行为具有多种表现形式, 本文仅对万维

网这种典型的应用模式进行了研究, 希望能引发后续工作对诸如分布式对等网络等其他应用层行为的关注.

现有实现网络服务质量保证的手段主要有增加网络资源与主动路由疏导等, 可视为“数量上”的改进. 这类局限于 Internet 单一体系结构下的改进必将使 Internet 的运动趋于某一物理极限而无法匹配迅猛增长的带宽需求; 因此添加免拥塞结构, 合理规避应用层行为对物理网络的影响, 实现“结构上”的改进成为了网络未来发展的新契机, 例如近期蓬勃发展的 CDN (content delivery network), P2P (peer to peer) 技术等. 本文研究成果为网络性能优化及管控研究提供新的角度, 同时指出基于均衡网络负载的结构改进方向: 1) 可规划设计物理拓扑引导用户行为逻辑分布, 从而均衡网络负载, 提升整体性能; 2) 在物理网络中通过散布镜像服务器或全局广播以减少网络中偏好性强的“富节点”的负载压力; 3) 为优化路由机理的设计提供理论支持; 4) 把对特定节点的恶意攻击视为一种虚拟网络行为, 则从用户行为整体特征的角度为网络安全性保证研究提供了思路.

-
- [1] Carlson J M, Doyle J 1999 *Phys. Rev. E* **60** 1412
 - [2] Willingers W, Paxson V 1998 *Notices of the American Mathematical Society* **45** 961
 - [3] Harchol-Balter M, Crovella M, Murta C 1999 *Journal of Parallel and Distributed Computing* **59** 204
 - [4] Tretyakov A Y, Takayasu H, Takayasu M 1998 *Physica A* **253** 315
 - [5] Ohira T, Sawatari R 1998 *Phys. Rev. E* **58** 193
 - [6] Fuks H, Lawniczak A T 1999 *Mathematics and Computers in Simulation* **51** 101
 - [7] Takayasu M, Fukuda K, Takayasu H 1999 *Physica A* **274** 140
 - [8] Takayasu M, Takayasu H, Fukuda K 2000 *Physica A* **277** 248
 - [9] Liu F, Shan X, Ren Y 2004 *Acta Phys. Sin.* **53** 273 (in Chinese) [刘 锋、山秀明、任 勇 2004 物理学报 **53** 273]
 - [10] Yuan J, Ren Y, Liu F, Shan X 2001 *Acta Phys. Sin.* **50** 1221 (in Chinese) [袁 坚、任 勇、刘 锋、山秀明 2001 物理学报 **50** 1221]
 - [11] Liu F, Ren Y, Shan X 2002 *Acta Phys. Sin.* **51** 1175 (in Chinese) [刘 锋、任 勇、山秀明 2002 物理学报 **51** 1175]
 - [12] Albert R, Barabási A 2002 *Reviews of Modern Physics* **74** 47
 - [13] Xia Y, Chi K 2005 *Physical Review E* **72** 026116
 - [14] Takayasu M, Takayasu H, Sato T 1996 *Physica A* **289** 595
 - [15] Sole R V, Valverde S 2001 *Physica A* **289** 595
 - [16] Li Y, Liu Y, Shan X, Ren Y 2005 *Chinese Physics* **14** 1280
 - [17] Mandelbrot B B 1982 *The Fractal Geometry of Nature* [W H Freeman & Co, San Francisco, USA] [Translated in Chinese in 1992]

Influence of virtual networks to internet collective behavior^{*}

Wang Lei Zhou Shu-Hua Yuan Jian Ren Yong Shan Xiu-Ming

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

(Received 30 April 2006 revised manuscript received 4 June 2006)

Abstract

A virtual network is defined as a logical network, in which users exhibit their access behaviors. Virtual networks rely on physical computer networks like the Internet, but have different topologies, and cause significant influence on the physical networks. In this paper, a novel two-tier model is proposed to study influences of virtual networks to Internet collective behavior. It is shown that the queue lengths of the node data packets present phase transition characteristics. Moreover, the phase transition critical point moves to the left and network performance is deteriorated. In a free flow, the nodes are independent of each other or short-range correlative. In the critical state, the nodes are long-range correlative, and there exists a higher power exponent H which means stronger long-range correlation. When the system state is on the right of critical point, virtual network behaviors make the network to present consistent long-range correlative characteristic.

Keywords: virtual network, cellular automation, phase transition, internet

PACC: 0540J, 8980J

^{*}Project supported by the National Natural Science Foundation of China (Grant Nos. 90204004, 90304005) and Tsinghua 985 Innovation Funding.