

无标度网络上名字无关的紧凑路由研究

唐明董^{1),2)} 刘建勋¹⁾ 张国清²⁾ 杨景³⁾

¹⁾(湖南科技大学计算机科学与工程学院 湖南 湘潭 411201)

²⁾(中国科学院计算技术研究所 北京 100190)

³⁾(中国移动通信研究院 北京 100053)

摘 要 紧凑路由是一类可扩展路由方法,能够在使用较短路径路由的同时,大幅降低节点路由表的空间开销.为解决 Internet 的路由扩展问题,无标度网络上的紧凑路由最近引起了关注.然而,以往的紧凑路由方法大多是名字有关的,即必须对网络节点重新命名或编址,这对于真实网络是不太实际的.针对无标度网络提出了一种名字无关的紧凑路由方法,即不需要显式改变节点的名字,任何节点只需要知道目的节点的原始名字就可以将包送达,因此具有更好的实用性.该方法在名字有关的紧凑路由方法基础上,引入一种名字映射系统.路由过程一般分为两个阶段:先由目的节点的原始名字查询其隐藏的地址,然后根据该地址路由.为了优化路由的性能指标,选择无标度网络中度最大的部分节点作为路由用的地标,并在这些地标上均匀且分布地为所有节点建立名字到位置的映射.分析和仿真表明,该路由方法可以在无标度网络上取得很高的路由性能,与以往最优的名字无关紧凑路由方法相比,在拉伸系数和路由表大小方面都有显著提高.

关键词 紧凑路由;路由方法;无标度网络;名字无关;拉伸系数

中图法分类号 TP393 **DOI号** 10.3724/SP.J.1016.2014.02353

Name-Independent Compact Routing in Scale-Free Networks

TANG Ming-Dong^{1),2)} LIU Jian-Xun¹⁾ ZHANG Guo-Qing²⁾ YANG Jing³⁾

¹⁾(School of Computer Science and Engineering, Hunan University of Science and Technology, Xiangtan, Hunan 411201)

²⁾(Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100190)

³⁾(China Mobile Research Institute, Beijing 100053)

Abstract Compact routing represents a category of scalable routing schemes that can greatly reduce routing table size of nodes while routing with nearly shortest paths. Aiming at addressing the scalability problem of Internet routing, compact routing on scale-free networks has received considerable attention. However, existing compact routing schemes are mostly name-dependent, i. e., nodes must be renamed to encode the network's topological information. Name-dependent routing schemes are thereby considered lacking of applicability in real networks. This paper proposes a name-independent compact routing scheme for scale-free networks, in which renaming node is not needed and packet delivery will be successful if only the source node knows the destination node's original name. The scheme incorporates a name mapping system based on a name-dependent compact routing scheme. Its routing process usually consists of two steps: querying the destination node's hidden address by its original one, then routing according to the address. To optimize the routing performance, the scheme uses a small number of high-degree

收稿日期:2013-07-19;最终修改稿收到日期:2014-03-18. 本课题得到国家自然科学基金(61100054,61272063,61370227)、湖南省自然科学基金(12JJ6064,12JJB009)、湖南省教育厅科学研究项目(11B048)资助. 唐明董,男,1978年生,博士,副教授,中国计算机学会(CCF)会员,主要研究方向为网络科学与技术. E-mail: tangmingdong@gmail.com. 刘建勋,男,1970年生,博士,教授,主要研究领域为服务计算、云计算. 张国清,男,1965年生,博士,研究员,主要研究领域为计算机网络、网络科学. 杨景,男,1952年生,博士,教授,主要研究领域为计算机网络、下一代互联网架构.

nodes as landmarks, and maps all nodes' original names to their addresses in an evenly distributed way on those landmarks. Comprehensive experiments show that the proposed scheme can achieve a significantly better performance in scale-free networks than the best existing name-independent compact routing scheme.

Keywords compact routing; routing schemes; scale-free networks; name-independent; stretch factor

1 引 言

路由是通信网络的一项基本功能. 一个路由方法的好坏主要由两个指标来衡量, 即节点路由表的大小和路径的质量. 但这两个指标通常是互相矛盾的. 传统的最短路径路由方法, 保证在任何两个节点之间路由都使用最优路径, 但是每个节点需要存放到所有其它节点的路由信息, 路由表项数为 $\Omega(n)$ (即至少与网络规模 n 成线性比例), n 代表网络规模. 当 n 很大时, 路由表规模将成为限制网络扩展的瓶颈. 以 Internet 为例, 域间路由系统是以选择最优路径为基础的, 因此核心路由表近年来急剧膨胀, 路由条目数目前已在 30 万以上^①, 引起的内存和计算开销大大加重了路由器的负担, 使当前的 Internet 面临严重的扩展问题^②. 即使路由器硬件技术能够跟上路由表的扩张速度, 由此产生的高能耗也将对网络扩展构成巨大挑战^[1]. 因此, 缩减路由表规模已经成为 Internet 可扩展路由的一个重要目标^[2-3].

紧凑路由(compact routing)为解决 Internet 的路由扩展问题提供了一种候选的方案^[4]. 紧凑路由是指产生的路由表规模为 $o(n)$ 的一类路由算法(即保证路由表规模随网络规模 n 呈亚线性增长), 能够大幅压缩路由表的空间开销. 然而压缩路由表可能会引起路径的拉伸, 即可能使用非最短路径路由, 从而降低了路由的效率. 为此, 各类紧凑路由研究都旨在尽可能同时优化路由表规模和路径的质量. 路径的好坏通常用拉伸系数(stretch factor)来衡量, 即实际路径相对于最短路径的拉伸程度. 任何紧凑路由方法的拉伸系数可以用 (a, b) 来表示, 指从任意节点 u 路由到另外一个节点 v 所使用的路径长度最多为 $ad(u, v) + b$, 其中 $d(u, v)$ 为 u 和 v 之间的最短路径长度. 如果 $b = 0$, 称路由方法具有倍数型拉伸系数 a . 显然, a 和 b 的值越小, 拉伸系数越优. 紧凑路由问题在一般的网络和特定结构的网络上都得到了广泛研究, 提出了许多通用或专用的紧凑路由方

法^[5]. 已有的紧凑路由方法根据节点是否根据网络拓扑编址, 而分为名字有关(name-dependent)的和名字无关(name-independent)的. 前者需要根据网络拓扑对节点重新命名(或编址), 节点的新名字需要嵌入拓扑信息, 以便指导路由; 后者指节点名字可以是任意的, 路由方法仅使用节点的原始名字就可以路由成功.

近年来的研究发现, 包括 Internet 在内的许多真实网络都是无标度网络, 即节点的连接度 k 满足幂律分布 $P(k) \sim k^{-\gamma}$ ($\gamma > 0$). 为了研究在 Internet 这类网络拓扑上的路由扩展性, 无标度网络上的紧凑路由问题最近得到关注^[6-11]. 然而以往针对无标度网络的紧凑路由方法大都是名字有关的. 名字有关的路由方法一般被认为在真实网络上不太实用, 原因是需要对节点进行重新命名, 并且拓扑的改变导致节点的名字也必须更新. 事实上, 大多数互联网路由专家认为, 为了可扩展性、灵活性和移动性的需要, 未来的互联网路由系统应该使用拓扑无关(亦即提供商无关)的地址^[4].

本文针对无标度网络提出了一种名字无关的、基于高度节点的地标路由方法——NIHDLR(Name-Independent Highest-Degree Landmark Routing). NIHDLR 使用的包首部长为 $O(\log n)$ ^③ 比特, 取得的拉伸系数为 $\min\{(3, 0), (2, d)\}$, 这意味着在任意两个节点之间路由, 路径长度不超过其最短路径的 3 倍以及最短路径的 2 倍加一个增量 d (d 是由网络拓扑结构决定的参数). 使用随机幂律图(random power-law graph)理论进行分析发现, d 可以取较小的常数, 且本文的紧凑路由方法产生的路由表平均大小不超过 $\tilde{O}(n^{1/2})$ 比特 ($\tilde{O}(f(n))$ 是 $O(f(n) \log^c n)$ 的缩写, c 为常数). 在无标度网络模型以及真实的 Internet 拓扑图上的仿真实验表明,

① BGP Routing Table Analysis Reports. <http://www.potaroo.com>, 2010, 1

② Report from the IAB workshop on routing and addressing. <http://tools.ietf.org/html/rfc4984>, 2007, 9

③ 本文中 \log 对数均是以 2 为底数

NIHDLR 的实际路径长度和路由表大小均要远低于通用的最优名字无关紧凑路由方法。

2 相关工作

自 20 世纪 90 年代以来,许多通用的紧凑路由方法被提出,大多数是名字有关的.其中,Cowen 方法^[12]和 TZ 方法^[13]是两种最具代表性的名字有关紧凑路由方法.Cowen 方法是第一个将拉伸系数降至(3,0)的通用紧凑路由方法,在每个节点上产生的路由表大小为 $\tilde{O}(n^{2/3})$ 比特.TZ 方法是 Cowen 方法的改进,在拉伸系数等于(3,0)的情况下,节点路由表大小降为 $\tilde{O}(n^{1/2})$ 比特.Cowen 方法和 TZ 方法均是基于地标的,即需要使用少数节点作为路由用的地标.地标路由算法中,往往只在节点路由表中存放到邻居和所有地标的路由信息,因此如果是以这些节点为目的,包可以沿最短路径交付,否则包将沿最短路径向离目的地最近的地标转发.如果到达该地标,将由它转交给目的节点.Cowen 方法和 TZ 方法的区别仅在于地标选择方法的不同.前者使用贪心算法来构造图的支配集,由此得到地标集;而后者则使用一种随机化的技术来选择地标.

进入 21 世纪以来,名字无关的紧凑路由方法开始引起关注.代表工作有 Arias 等人^[14]提出的路由方法和 Abraham 等人^[15]提出的路由方法.Arias 方法在任意的网络上可以保证拉伸系数不超过(5,0),路由表大小不超过 $\tilde{O}(n^{1/2})$ 比特.Abraham 方法产生的拉伸系数和路由表大小分别为(3,0)和 $\tilde{O}(n^{1/2})$ 比特,与 TZ 方法取得了近乎相同的理论界限.由于 Gavoille 和 Gengler 证明不存在路由方法在任意网络上使倍数型拉伸系数低于 3 且每个节点的路由表大小为 $o(n)$ ^[16].因此 Abraham 方法和 TZ 方法一样都是近乎最优的通用紧凑路由方法.上述紧凑路由方法都是针对任意拓扑结构网络的.然而,真实的网络往往呈现特定的拓扑结构特征,针对特定拓扑结构定制的紧凑路由方法往往比通用的方法具有更好的性能,包括树(tree)、平面图(planar graph)等^[17-19].现实中,许多网络(包括互联网、WWW、非结构化 P2P 网络等)都是无标度网络^[20-25].这类网络中,多数节点度很小,而少数节点度很大.这些大度节点在网络连通性和通信方面常常扮演了十分关键的角色.无标度网络(或称幂律网络)上的紧凑路由由算法最近引起了许多关注.对它们的综述如下.

文献[6]分析和仿真了 TZ 方法在幂律网络上的性能,发现在产生很小路由表的同时平均拉伸度仅为 1.1 左右,远低于理论上的拉伸系数(3,0).文献[10]提出了一种贪心的地标选择方法,通过实验发现在幂律网络上实际的倍数型拉伸系数不超过 2.文献[11]针对无限变异幂律图(infinite-variance power-law random graph)提出了改进的紧凑路由方法.文献[26]提出使用 k -核分解的办法来选择 Internet 拓扑图中的核心节点作为地标,可以进一步提高紧凑路由算法在 Internet 拓扑图上的性能.文献[7-8]提出了具有增量型拉伸系数(1, d)的名字有关紧凑路由方法,路由表大小上限为 $O(e \log^2 n)$ 比特,实验发现在幂律网络上 d 和 e 都可以取较小的值.上述工作都没有在幂律网络上针对路由指标给出严格的理论界限.文献[9,27]使用随机幂律图理论分析了提出的紧凑路由方法的性能,证明在随机幂律图上路由表的大小具有比一般网络上低得多的上限,而拉伸系数仍为(3,0).文献[28]提出在稀疏的网络上(包括幂律网络)可以进一步将拉伸系数降至(2,0),代价是转发时间增加至 $\tilde{O}(n^{1/2})$.

以上针对无标度网络的紧凑路由方法都是名字有关的.针对无标度网络的名字无关的紧凑路由方法尚不多见.文献[29]提出了一种名字无关的紧凑路由方法,具有拉伸系数(2,2 d),其中 d 在随机幂律图上可以取常数,路由表平均大小不超过 $\tilde{O}(n^{1/2})$ 比特.本文的工作通过在路由时引入握手机制(handshaking)不仅进一步优化了路由的拉伸系数上限,而且实际的路径拉伸度比^[29]也有显著的降低.

3 NIHDLR 方法

3.1 原理

名字无关的紧凑路由方法一般是在名字有关的紧凑路由方法之上加上一个节点名字/地址映射系统.除非特别说明,本文中节点名是指节点固有的与网络拓扑无关的名字,节点地址是指重新赋予的与拓扑有关的名字.路由过程一般包含两部分:先利用名字/地址映射系统根据目的节点名查询其地址,然后利用名字有关的路由算法根据目的地址将包转发至目的节点.

与名字有关路由方法相比,名字无关方法的设计更加复杂,表现在:名字无关的方法在路由时增加了查询目的节点地址的过程,而维护节点名字到地

址的映射则增加了节点的内存开销. 一般来说, 这里存在一个折中(trade-off)问题. 如果每个节点都存储整个网络所有节点的名字/地址映射, 那么源节点在本地就可以查询到任意目的节点的地址, 不会引起额外的路径开销, 但是此时每个节点需要增加 $\tilde{O}(n)$ 比特的存储空间; 另一方面, 如果将映射表分割成 n 块均匀地分散到 n 个节点上存放, 那么节点的存储开销基本上没有多少增加, 但是查询目的节点地址可能需要很长的路径. 上述这两种极端情况都是不可取的. 映射系统的设计必须保证既不能在某些节点上占用太多的空间又必须保证查询目的地址不能引入太多的路径代价.

为了平衡映射系统引起的查询路径长度和空间开销, Abraham 方法^[15]将节点名字/地址映射表分成 \sqrt{n} 份, 每个节点维护一份, 保证每个节点附近的 $\tilde{O}(n^{1/2})$ 个邻居拥有一份完整的节点名字/地址映射. 这样保证了较小范围内就可以查询到任何目的节点的地址. Abraham 方法在任意拓扑类型的网络上产生的包首部长为 $O(\log^2 n / \log \log n)$ 比特, 拉伸系数为 $(3, 0)$, 节点的内存开销为 $\tilde{O}(n^{1/2})$ 比特.

NIHDLR 方法利用了无标度网络的结构特征. 在无标度网络上, 少数大度节点扮演网络和路由 hub 的角色, 多数节点与这些大度节点所形成的核心(core)直接相连或只有很少的跳数; 而且大度节点之间往往也具有很小的距离^[30-31]. 在这种情况下, 如果把名字/地址映射表均匀地分散存放到大度节点上, 而在其它节点不需要存放, 直觉上, 查询目的地址引起的路径开销不会太大. 并且, 这样做仅仅会引起少数高度节点的内存开销增大. 而如果选择的高度节点足够多, 比如为 \sqrt{n} , 那么高度节点上增加的空间开销也会很小.

NIHDLR 方法在底层使用的名字有关的紧凑路由方法, 是改编自经典的 TZ 方法^[13], 称为 TZ' 方法. 不同于 TZ 方法, TZ' 方法选择网络中的大度节点作为地标. 下面对 TZ' 方法进行介绍.

3.2 一种名字有关的紧凑路由方法

设网络用无向无权图 $G=(V, E)$ 表示, V 是节点集合, E 是边的集合. 设 $|V|=n, |E|=m, L \subseteq V$ 表示地标集合. 对任意节点 $u \in V$, 定义节点 u 的 ball 为 $B(u) = \{v \in V \mid d(u, v) < d(u, L)\}$, 其中 $d(u, L)$ 代表 u 到所有地标的最小距离. 定义节点 u 的 cluster 为 $C(u) = \{v \in V \mid u \in B(v)\}$. 下面以图 1 为例进行说明. 假设节点 a, b, c, d 是地标, 那么节点 u, v, w, s, t 的 ball 和 cluster 如表 1 所示.

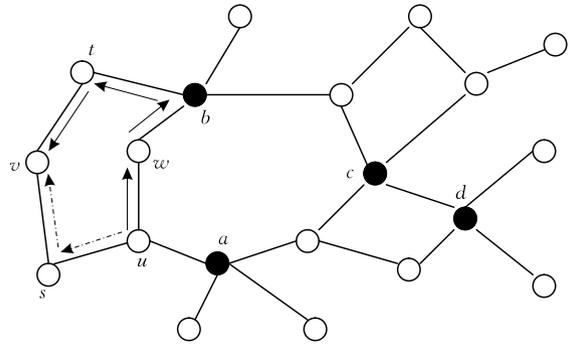


图 1 说明 TZ' 方法的例子(填充的节点 a, b, c, d 代表地标. 从节点 u 发送一个包至节点 v , 实线箭头代表使用的路径, 虚线箭头代表最短路径)

表 1 图 1 中各节点的 ball 和 cluster 计算

	s	t	u	v	w
$B()$	$\{u, v\}$	\emptyset	\emptyset	$\{s, t\}$	\emptyset
$C()$	$\{v\}$	$\{v\}$	$\{s\}$	$\{s\}$	\emptyset

TZ' 方法的设计分预处理部分和转发算法部分来介绍. 预处理部分包含如下过程:

(1) 地标集构造. 选择 G 中前 n^x 个度数最大的节点作为地标集 L , 其中 $0 < x < 1$ 是可以由用户设置的参数.

(2) 地址分配. 对每个非地标节点 $u \in V - L$, 分配地址 $(u, L(u), e_{L(u)}(u))$, 记为 $Address_{TZ'}(u)$. 其中 $L(u)$ 是离 u 最近的那个地标的名字, 而 $e_{L(u)}(u)$ 代表 $L(u)$ 到 u 的最短路径上与 $L(u)$ 相连的端口. 对每个地标, 就用它的名字作为地址.

(3) 路由表设置. 对每个非地标节点 $u \in V - L$, 它的路由表设置为

$Table_{TZ'}(u) = \{(v, e_u(v)) \mid v \in L \cup C(u) \cup B(u)\}$, 即存储到 $L, C(u)$ 和 $B(u)$ 中每个节点的最短路由信息, 其中 $e_u(v)$ 代表 u 的与它到 v 的最短路径相连的端口. 而对每个 $l \in L$, 其路由表只需要包含到其它地标的路由信息, 即 $Table_{TZ'}(l) = \{(v, e_l(v)) \mid v \in L - \{l\}\}$.

值得说明的是, TZ 方法在每个非地标节点 u 上产生的路由表为 $Table_{TZ}(u) = \{(v, e_u(v)) \mid v \in L \cup C(u)\}$. 因此与 TZ 方法相比, TZ' 方法增加了到 $B(u)$ 的最短路由信息. 它的作用是降低 NIHDLR 方法的拉伸系数.

以图 1 为例, 节点 u 的路由表只包含到 $\{a, b, c, d, s\}$ 的路由信息, 节点 v 的路由表只包含到 $\{a, b, c, d, s, t\}$ 的路由信息. 考虑到 TZ' 方法与 TZ 方法的地址分配过程, 路由表构造过程采用的算法基本一致, 这些过程的详细设计这里不再介绍. TZ' 方法的转

发算法与 TZ 方法也基本相同. 设包首部含有目的节点 v 的地址 $(v, L(v), e_{L(v)}(v))$, 在收到包的每个节点 u 上执行如下转发操作:

- (1) 如果 $u=v$, 那么包到达终点.
- (2) 否则, 如果 $u=L(v)$, 那么使用本地端口 $e_{L(v)}(v)$ 转发包.
- (3) 否则, 如果 $Table_{TZ'}(u)$ 中有到 v 的路由信息, 那么使用 $e_u(v)$ 转发包.
- (4) 否则, 在 $Table_{TZ'}(u)$ 中查找到 $L(v)$ 的路由信息, 用 $e_u(L(v))$ 转发包.

以图 1 为例, 从节点 u 发送一个包到 v , 根据上面的转发算法, 使用的路径将是 $u-w-b-t-v$. 但是 u 和 v 之间的最短路径是 $u-s-v$. 因此, 同其它紧凑路由方法一样, TZ' 方法也可能使用非最短路径路由.

3.3 基本设计

分预处理和转发算法两部分来介绍 NIHDLR 方法的设计. 建立在 TZ' 方法的基础上, NIHDLR 方法的预处理过程描述如下.

(1) 像 TZ' 方法那样选择连接度最大的那些节点作为地标. 取 $x=1/2$, 即使用 G 的前 \sqrt{n} 个最大度节点构造地标集 L .

(2) 像 TZ' 方法那样对每个节点分配地址和构造路由表. 对节点 u 分配的地址和构造的路由表仍然分别用 $Address_{TZ'}(u)$, $Table_{TZ'}(u)$ 表示.

(3) 构造节点名字/地址映射表, 将映射表分布地存放在所有地标上, 使每个节点都能够查询任意目的节点的地址.

预处理过程的关键是如何建立节点名字/地址映射. 我们设计了一种 Hash 函数 H , 它将一个节点 v 映射到一个地标, 记为 $H(v)$. 每个节点都可以为任意目的节点 v 计算 $H(v)$. 在地标 $H(v)$ 上存放 v 的名字和地址对, 即 $(v, Address_{TZ'}(v))$. 为了使每个地标存放的名字/地址对数量比较均匀, 我们让每个普通节点随机选择一个地标. 根据 Chernoff bound 定理^[32], 当地标数量为 \sqrt{n} 时, 这样做可以保证每个地标存储的名字/标记数在 $\tilde{O}(n^{1/2})$ 中. 对 Hash 函数设计的描述如下. 设每个节点都知道所有地标的名字, 可以按地标名的字母顺序使用 $0, 1, 2, \dots, \sqrt{n}-1$ 等数字对地标编号. 因此 Hash 函数 H 可以使用 $[0, \sqrt{n}-1]$ 中的数字作为输出值, 每个数字可以唯一确定一个地标. 因此, 如果使用 $\{0, 2, \dots, n-1\}$ 中的数字作为节点名字, Hash 函数 H 的设计将十分简单. 设节点名字用 $\log n$ 比特编码, 那么从

节点名字中抽取 $(\log n)/2$ 位作为输出结果即可, 如节点名字的低 $(\log n)/2$ 位. 为了便于理解, 图 2 给出了一个当 $n=16$ 时 H 的设计举例. 设节点名字取自于 $\{0, 1, \dots, n-1\}$, 设 $4(4=\sqrt{16})$ 个地标节点的名字是 $\{3, 8, 11, 15\}$. 如果取节点名字的低 $(\log n)/2$ 位作为 Hash 结果, 那么如图 2 中所示, 节点 $\{1, 6, 7, 9, 12\}$ 的 Hash 结果将分别是 $\{1, 2, 3, 1, 0\}$. 根据地标名字的大小顺序, 输出值 $0, 1, 2, 3$ 将分别对应地标 $3, 8, 11, 15$. 这就将每个节点和地标对应了起来. 如果节点可以使用任意的字符串作为名字, Hash 函数 H 的设计可以借鉴文献[15]中的随机化构造策略, 同样可以在常数时间内完成映射, 且保证最多将 $\tilde{O}(n^{1/2})$ 个节点映射到同一个地标上. 具体的构造过程参考文献[15], 这里不再赘述.

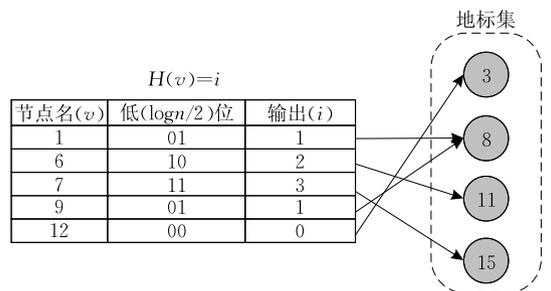


图 2 Hash 函数 H 的设计举例(设节点名字取自于 $\{0, 1, \dots, n-1\}$, $n=16$)

设源节点为 u , 目的节点为 v , NIHDLR 方法在一个节点上转发包的决策过程如下:

(1) 如果 v 在 u 的路由表中, 那么 u 可以将包沿最短路径发送到 v .

(2) 否则, u 计算 $H(v)$, 确定存放 v 的位置信息的地标.

(3) u 将包发送给 $H(v)$. 根据路由表构造策略, $H(v)$ 在 u 的路由表中, 因此包可以沿最短路径路由到 $H(v)$. 注意, 在此过程中, 如果中间节点 w 的路由表中包含到 v 的路由信息, 那么 w 将沿到 v 的最短路径来转发包. 因此可能使用一条不经过 $H(v)$ 的捷径来交付包.

(4) 当包到达 $H(v)$ 后, $H(v)$ 查询 v 的地址 $(v, L(v), e_{L(v)}(v))$, 用它重写包首部, 然后按照 TZ' 方法的转发算法将包发送至目的节点 v .

以上的路由过程在最坏情况下所使用的路径如图 3 中的实线箭头所示(为了清楚起见, 没有画出节点之间的连接). 此时的路径长度为

$d_R(u, v) = d(u, L(u)) + d(L(u), H(v)) + d(L(v), v)$. 下面提出对 NIHDLR 的路径质量进行优化的方法.

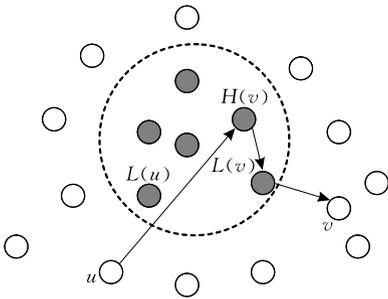


图 3 NIHDLR 方法的路由过程的说明(带阴影的节点代表地标, $L(v)$ 代表离节点 v 最近的地标, $H(v)$ 代表存储节点 v 标记的地标)

3.4 优化

为了进一步优化路径质量,我们为 NIHDLR 引入了握手机制(handshaking):当以 v 为目的地的包到达 $L(v)$ 时, $L(v)$ 将查询 v 的地址 $(v, L(v))$, $e_{L(v)}(v)$, 并将其封装成一个消息返回给源节点 u . 返回该消息可以有多种路由方式. 例如, $L(v)$ 可以先将消息发送给 $H(u)$, 然后由 $H(u)$ 根据 u 的地址把封装目的地地址的消息发送给 u . 这样就完成了一次握手. 完成握手后, u 在与 v 通信时可以直接沿最短路径将包发送给 $L(v)$, 然后由 $L(v)$ 转交给 v . 这个过程可以由图 4 中的实线箭头来描述. 可见, 握手后最坏情况下的路径长度为

$$d_R(u, v) = d(u, L(v)) + d(L(v), v).$$

后面我们将证明最坏情况下的拉伸系数.

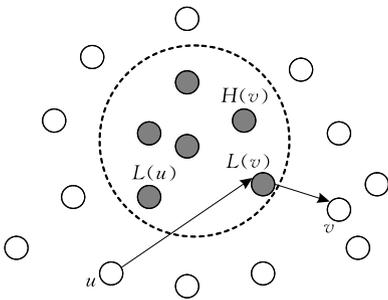


图 4 握手后的路由过程的说明(带阴影的节点代表地标, $L(v)$ 代表离节点 v 最近的地标, $H(v)$ 代表存储节点 v 标记的地标)

4 性能分析

下面分析 NIHDLR 方法的包首部长度的、拉伸系数、路由表大小、预处理时间和转发时间等指标的性能.

4.1 包首部长度的

NIHDLR 方法产生的包首部长度的为 $O(\log n)$ 比

特,理由如下:对任意节点 v , 它的地址为 $(v, L(v))$, $e_{L(v)}(v)$, 由于节点名、端口号都可用 $O(\log n)$ 比特表示, 故地址长度为 $O(\log n)$ 比特. 包首部一开始只包含目的节点的名字, 后来用目的节点的地址重写, 不论哪一种情况, 包首部长度的都在 $O(\log n)$ 比特中. 而 Abraham 方法产生的包首部长度的为 $O(\log^2 n / \log \log n)$ 比特, 因此 NIHDLR 方法的在包首部产生的开销方面要优于 Abraham 方法.

4.2 拉伸系数

这里对 NIHDLR 方法分别在握手机制和具有握手机制的情况下, 产生的拉伸系数进行分析.

定理 1. 不考虑握手机制时, NIHDLR 方法在任意图 G 上具有不超过 $(2, 2d)$ 的拉伸系数, 其中 $d = \max\{d(u, v) \mid u, v \in L \wedge u \neq v\}$, 即图 G 中所有不同地标对之间的最大距离.

证明. 设源节点 u 发送一个包给目的节点 v . 如果 $v \in L \cup C(u) \cup B(u)$ 那么可以直接使用 v 的名字沿最短路径路由到 v . 否则, 路由过程如图 3 所示. 因此, 从 u 到 v 的路径长度 $d_R(u, v)$ 一定满足

$$d_R(u, v) \leq d(u, H(v)) + d(H(v), L(v)) + d(L(v), v).$$

根据三角形不等式原理, 显然有 $d(u, H(v)) \leq d(u, L(u)) + d(L(u), H(v))$, 所以

$$d_R(u, v) \leq d(u, L(u)) + d(L(u), H(v)) + d(H(v), L(v)) + d(L(v), v).$$

因为 v 不在 $Table(u)$ 中, 所以 $d(u, v) > d(u, L(u))$ 且 $d(u, v) > d(L(v), v)$ 成立, 故

$$d_R(u, v) \leq 2d(u, v) + d(L(u), H(v)) + d(H(v), L(v)).$$

又因为 $d(L(u), H(v)) \leq d$, $d(H(v), L(v)) \leq d$, 所以

$$d_R(u, v) \leq 2d(u, v) + 2d.$$

因此 NIHDLR 方法的拉伸系数等于 $(2, 2d)$.

证毕.

定理 2. 考虑握手机制, 握手后 NIHDLR 方法在任意图 G 上具有 $\min\{(3, 0), (2, d)\}$ 的拉伸系数, d 的定义如定理 1.

证明. 握手后, 从源节点 u 发送一个包给目的节点 v , 最坏情况下的路由过程如图 4 所示. 因此, 从 u 到 v 的路径长度 $d_R(u, v)$ 一定满足

$$d_R(u, v) \leq d(u, L(v)) + d(L(v), v).$$

根据三角形不等式原理, 显然有 $d_R(u, L(v)) \leq d(u, v) + d(L(v), v)$, 所以

$$d_R(u, v) \leq d(u, v) + 2d(L(v), v).$$

因为 $v \notin B(u) \cup C(u)$, 所以 $d(u, v) > d(L(v), v)$, 故有 $d_R(u, v) \leq 3d(u, v)$.

下面证明 $d_R(u, v) \leq 2d(u, v) + d$. 根据三角形不等式原理, 亦有 $d(u, L(v)) \leq d(u, L(u)) + d(L(u), L(v))$. 故有

$d_R(u, v) \leq d(u, L(u)) + d(L(u), L(v)) + d(L(v), v)$. 再次考虑到 $d(u, v) > d(u, L(u))$ 且 $d(u, v) > d(L(v), v)$, 所以有

$d_R(u, v) \leq 2d(u, v) + d(L(u), L(v)) \leq 2d(u, v) + d$. 结合 $d_R(u, v) \leq 3d(u, v)$ 和 $d_R(u, v) \leq 2d(u, v) + d$, 可以得出 NIHDLR 方法具有最大拉伸系数 $\min\{(3, 0), (2, d)\}$. 证毕.

在无标度网络上 d 到底能取什么样的值, 我们将通过 Lu^[30] 和 Chung 等人^[31] 提出的随机幂律图模型及相关理论来分析. 假设幂指数的取值在 $2 < \gamma < 3$ 中, 事实上多数无标度网络都是满足该假设的. 下面给出随机幂律图的定义^[30-31].

定义 1. 设 $G(\mathbf{w})$ 表示由度序列 $\mathbf{w} = (\omega_1, \omega_2, \dots, \omega_n)$ 产生的随机图集合, 其中 ω_i 表示节点 v_i 的期望度, 任意两个节点 v_i, v_j 之间存在一条边的概率等于 $\omega_i \omega_j / \sum_{k=1}^n \omega_k$. 如果 $\mathbf{w} = (\omega_1, \omega_2, \dots, \omega_n)$ 满足幂律分布, 即 $P(\omega_i) \sim \omega_i^{-\gamma}$, 那么称 $G \in G(\mathbf{w})$ 为随机幂律图^[30-31].

设 k -Core 为 $G(\mathbf{w})$ 中期望度大于或等于 k 的那些节点集合. 那么它导出的子图的直径具有以下性质.

引理 1. 设 G 是 $G(\mathbf{w})$ 的一个采样, 给定 $t \geq n^\delta$ ($0 < \delta < \frac{3-\gamma}{\gamma-2}$, 可任取), $G(\mathbf{w})$ 的 t -Core 导出的子图有 $1-n^{-2}$ 的概率直径不超过 $(1+o(1)) \frac{\log n}{(3-\gamma)\log t}$.

证明. 请参考文献[30]第 62 页的 Lemma 12 和文献[31]的 Claim 4. 1. 过程略. 证毕.

引理 1 意味着当 n 很大时, 由随机幂律图中的少数最大度节点导出的子图具有的直径趋近于常数. 该子图可以看成是幂律图的核(Core). 于是我们可以得到以下的推论.

推论 1(引理 1 的推论). 设 G 是 $G(\mathbf{w})$ 的一个采样, 地标集 L 由 $G(\mathbf{w})$ 的前 n^x ($0 < x < 1$) 个期望度最大的节点构成, 那么 d 有 $1-n^{-2}$ 的概率具有常数上限.

证明. 设 $G(\mathbf{w})$ 的期望度分布满足 $P(k) \sim k^{-\gamma}$, 从而可得累积度分布公式为 $P(z \geq k) = ck^{1-\gamma}$, 其中 c 可取常数. 现在计算 L 中节点的最小期望度 τ . 根据累积度分布公式有 $c\tau^{1-\gamma} \approx n^x/n$, 因此有

$$\begin{aligned} c\tau^{1-\gamma} &\approx n^x/n \Rightarrow \log c - (\gamma-1)\log \tau \\ &\approx (x-1)\log n \Rightarrow \log \tau \\ &\approx (1-x)\log n / (\gamma-1) + \log c / (\gamma-1) \Rightarrow \tau \\ &\approx c_1 n^{(1-x)/(\gamma-1)}, \end{aligned}$$

其中 c_1 是依赖于 γ 和 c 的常数.

L 可以看成是 $G(\mathbf{w})$ 的 τ -Core. 根据引理 1, L 导出的子图的直径有 $1-n^{-2}$ 的概率不超过

$$(1+o(1)) \frac{\log n}{(3-\gamma)\log \tau} = (1+o(1)) \frac{(\gamma-1)\log n}{(3-\gamma)((1-x)\log n + \log c)},$$

由于 γ, c 和 x 都可以取常数, 这个表达式具有常数上限. 根据 d 的定义, d 将小于或等于 L 导出的子图的直径. 因此 d 具有常数上限. 证毕.

4.3 路由表大小

NIHDLR 方法产生的路由表大小分地标和非地标节点两种情况分析. 对每个地标 l , 它的存储开销包含两部分. 一部分是到所有其它地标的路由信息, 即 $\{(v, e_l(v)) \mid v \in L - \{l\}\}$, 另一部分是节点名字/地址映射信息, 即 $\{(v, Address_{TZ'}(v)) \mid v \in V \wedge H(v) = l\}$. 当地标数量取 \sqrt{n} 时, 每个地标存放的节点名字/地址映射信息为 $\tilde{O}(n^{1/2})$ 比特, 因此每个地标的路由表大小不超过 $\tilde{O}(n^{1/2})$ 比特.

每个非地标节点的路由表设置为 $Table(u) = Table_{TZ'}(u) = \{(v, e_u(v)) \mid v \in L \cup C(u) \cup B(u)\}$. 非地标节点不需要存储节点名字/地址映射信息. 因此要求出路由表大小的界限, 关键是计算出 $B(u)$ 和 $C(u)$ 的大小. 为此, 我们仍然使用了随机幂律图理论来进行分析.

引理 2. 设 $\beta = \frac{\gamma-2}{2\gamma-3} + \epsilon$, $\beta' = \frac{1-\beta}{\gamma-1}$, $\alpha = \beta'(\gamma-2) + \frac{(2\gamma-3)\epsilon}{\gamma-1}$, 其中 ϵ 是满足 $n^{\frac{(2\gamma-3)\epsilon}{\gamma-1}} \geq \frac{2(\gamma-1)}{\gamma-2} \ln n$ 的任意正实数(显然 $\epsilon = o(1)$, 即当 n 足够大时, ϵ 将非常小). 设 G 是 $G(\mathbf{w})$ 的一个采样, 对所有 $u \in V(G)$, 至少有 $1-3/n^{-2}$ 的概率满足 $|B(u)| = |\{u' \in V(G); d(u, u') < d(u, \tau\text{-Core}(\mathbf{w}))\}| = O(n^\alpha)$,

其中 $\tau = n^{\beta'}$.

证明. 请参考文献[9]的第 9 页, Lemma 7. 具体过程略. 证毕.

推论 2(引理 2 的推论). 设 G 是 $G(\mathbf{w})$ 的一个采样, L 是地标集, 即 \mathbf{w} 的前 n^x ($0 < x < 1$) 个期望度最大的节点, 对所有 $u \in V(G)$, 至少有 $1-3/n^{-2}$ 的概

率满足 $|B(u)| = O(n^\alpha)$, 其中 $\alpha = \frac{(1-x)(\gamma-2)}{\gamma-1} + \xi$, ξ 足够小.

证明. 根据推论 1 已知, L 的最小期望度 $\tau = c_1 n^{(1-x)/(\gamma-1)}$, 进一步可转换为 $\tau = n^{\frac{1-x}{\gamma-1} + \epsilon'}$, 其中 $\epsilon' = o(1)$ 足够小. 根据引理 2 可知: $\alpha = \left(\frac{1-x}{\gamma-1} + \epsilon'\right)(\gamma-2) + \frac{(2\gamma-3)\epsilon}{\gamma-1}$, 即 $\alpha = \frac{(1-x)(\gamma-2)}{\gamma-1} + \xi$, 其中 $\xi = (\gamma-2)\epsilon' + \frac{(2\gamma-3)\epsilon}{\gamma-1}$. 因为 ϵ' 和 ϵ 都可以取足够小的值, 所以 ξ 也可以取很小的值. 证毕.

对所有 $u \in V(G)$, 我们通过推论 2 给出了 $|B(u)|$ 的理论上限. 遗憾的是 $|C(u)|$ 无法像 $|B(u)|$ 那样通过从 u 开始的邻域扩张法 (neighborhood expansion) 来统计分析它的界限. $C(u)$ 中包含哪些节点严格依赖于 u 及其邻居的距离信息, 而这些必须在具体的图产生后才能计算出. 然而, 我们可以计算出对全部 $u \in V(G)$, $C(u)$ 的平均大小的上限.

引理 3. 对任意图 G , 有

$$\sum_{u \in V(G)} |B(u)| = \sum_{u \in V(G)} |C(u)|,$$

即 $|C(u)|$ 和 $|B(u)|$ 具有相同的平均大小.

证明. 根据 $C(u)$ 和 $B(u)$ 的定义, 对任意 $v \in B(u)$, 必有 $u \in C(v)$, 反过来亦成立. 由于这样的 u, v 总是成对出现的, 所以引理 3 成立. 证毕.

根据以上的分析, 我们可以得出以下的重要结论.

定理 3. 设 G 是 $G(w)$ 的一个采样, 设 L 是地标集, 即 w 的前 n^x ($0 < x < 1$) 个期望度最大的节点, 对所有非地标节点 $u \in V(G)$, 平均的路由表大小至少有 $1 - 3/n^{-2}$ 的概率满足 $|\overline{Table(u)}| = \tilde{O}(n^x + n^\alpha)$, 其中 $\alpha = \frac{(1-x)(\gamma-2)}{\gamma-1} + \xi$, ξ 足够小.

证明. 因为 $|L| = O(n^x)$, $|B(u)| = O(n^\alpha)$ 且 $|\overline{C(u)}| = |\overline{B(u)}| = O(n^\alpha)$, 而每个路由表项可以用 $O(\log n)$ 编码, 所以很容易推出 $|\overline{Table(u)}| = \tilde{O}(n^x + n^\alpha)$. 证毕.

推论 3 (定理 3 的推论). NIHDLR 方法可以使平均路由表大小 $|\overline{Table(u)}|$ 在随机幂律图上具有上限 $\tilde{O}(\sqrt{n})$ 比特.

证明. 因为 $2 < \gamma < 3$, 故 $\alpha = \frac{(1-x)(\gamma-2)}{\gamma-1} + \xi \leq (1-x)/2 + \xi$, 所以非地标节点的平均路由表大

小为 $|\overline{Table(u)}| = \tilde{O}(n^x + n^{(1-x)/2 + \xi})$. 当 $x = 1/2$ 时, $|\overline{Table(u)}|$ 具有上限 $\tilde{O}(\sqrt{n})$. 而地标节点的路由表大小也具有上限 $\tilde{O}(\sqrt{n})$, 最后可以得出全部节点的平均路由表大小一定在 $\tilde{O}(\sqrt{n})$ 中. 证毕.

4.4 预处理时间与转发时间

定理 4. 设 G 是 $G(w)$ 的一个采样, 设 L 是地标集, 即 w 的前 n^x ($0 < x < 1$) 个期望度最大的节点, 使用 NIHDLR 路由方法, 预处理时间至少有 $1 - 3/n^{-2}$ 的概率为 $O(mn^x + n^{1+\alpha} \log n)$, 其中 $\alpha = \frac{(1-x)(\gamma-2)}{\gamma-1} + \xi$, ξ 足够小. 在每个节点上的路由转发时间为 $O(1)$.

假设使用集中式的算法进行预处理. 为了选择度最大的前 \sqrt{n} 个最大度节点构造地标集, 可以用一次宽度优先搜索算法搜索整个网络 G , 然后再对所有节点按度数排序, 因此需要的时间为 $O(m + n \log n)$. 用 TZ' 方法对节点 u 编址的时间取决于搜索其最近的地标所需要的时间, 实际上也等于计算 $B(u)$ 所需要的时间. 在随机幂律图上计算 $B(u)$ 所需要的时间 $T(B(u))$ 如引理 4 所示.

引理 4 [9]. 设 G 是 $G(w)$ 的一个采样, 使用 NIHDLR 路由方法, 对所有 $u \in V(G)$, $T(B(u))$ 至少有 $1 - 3/n^{-2}$ 的概率满足 $T(B(u)) = \tilde{O}(n^\alpha)$, 其中 $\alpha = \frac{(1-x)(\gamma-2)}{\gamma-1} + \xi$, ξ 足够小.

证明. 根据参考文献 [9] 的第 9 页, Lemma 7, 可以知道 $B(u)$ 的内部连接数 $E(B(u))$ 有 $1 - 3/n^{-2}$ 的概率在 $O(n^\alpha \log n)$ 中. 而使用宽度优先搜索 $B(u)$ 所需要的时间为 $O(E(B(u)))$, 所以得证. 证毕.

用 TZ' 方法对节点 u 创建路由表, $B(u)$ 和 $C(u)$ 部分所需要的时间等价于计算 $B(u)$ 的时间, 因此也是 $O(n^\alpha \log n)$. 可见, 为所有节点编址和创建到 $B(u)$ 和 $C(u)$ 部分路由表项的时间开销为 $O(n^{1+\alpha} \log n)$. 为了完成为所有节点创建到 L 的路由表项, 可从 L 中的每个地标发起宽度优先遍历, 时间开销为 $O(mn^x)$.

以上是使用 TZ' 方法进行预处理的时间开销. 下面分析创建节点名字/地址映射表的时间开销. 这里的主要开销来自于为每个节点计算 Hash 函数 H . 由于 H 的计算时间复杂度可以是常数 $O(1)$, 所以创建节点名字/地址映射表的时间开销应当为 $O(n)$. 综上所述, NIHDLR 方法的总的预处理时间应当为

$$O(m+n\log n)+O(n^{1+\alpha}\log n)+O(mn^x)+O(n),$$

即 $O(mn^x+n^{1+\alpha}\log n)$, 其中 $\alpha=\frac{(1-x)(\gamma-2)}{\gamma-1}+\xi$,

ξ 足够小. 定理 4 得证.

推论 4(定理 4 的推论). NIHDLR 方法的预处理时间可以控制在 $O(mn^{1/2})$ 中.

证明. 因为 $2<\gamma<3$, 故 $\alpha=\frac{(1-x)(\gamma-2)}{\gamma-1}+$

$\xi\leq(1-x)/2+\xi$, 所以 NIHDLR 方法的预处理时间为 $O(mn^x+n^{(3-x)/2+\xi}\log n)$. 当 $x=1/2$ 时, 可得该预处理时间的上限 $O(mn^{1/2})$. 证毕.

如同以往的工作一样^[12-14], 我们假设路由表项是用标准 Hash 表^[33]的方式来存储, 可以保证查询通往目的节点的端口号可以在常数时间内完成. 类似地, 根据目的节点名查询存放其地址的地标, 以及在地标上查询目的节点的地址也是使用 Hash 表实现. 故在路由过程中, 每个节点的决策和转发时间也是常数.

5 仿 真

为了验证提出的方法在无标度网络上的实际性能, 我们在随机幂律图模型和真实的 Internet AS (Autonomous System, 自治系统) 图上开展了仿真实验. 我们使用不同的幂指数 γ 和网络规模 n 来产生随机幂律图, 其中 γ 取 $\{2.1, 2.2, 2.3, 2.4, 2.5\}$ 中的值, n 取 $\{1000, 2000, \dots, 10000\}$ 中的值. 我们为每对 γ 和 n 生成 10 个随机幂律图, 仿真结果是在 10 个随机幂律图上进行平均所得. Internet AS 图使用 CAIDA^① 的 Skitter 项目测量的 AS 拓扑数据 ($\gamma\approx 2.25$), 其中节点数 $n=9204$, 连接数 $m=28959$. 仿真中设置地标集的大小为 $|L|=\sqrt{n}$. 仿真测试了 NIHDLR 的平均路由表大小, 平均拉伸系数等指标. 其中平均拉伸系数等于所有不同节点对之间的实际路径长度与最短路径长度之比的平均值. 我们还将 NIHDLR 与最优的通用名字无关紧凑路由方法——Abraham 方法进行了比较.

5.1 拉伸系数

图 5 给出了随机幂律图上 d 的测试结果. 可以看到, d 随网络规模的变化大致保持稳定(波动源于不同规模的随机幂律图的连接密度有差异), 这与我们的分析结果基本相符. 对于 $\gamma=2.1$ 的情况, d 的平均值稍大于 4. 而随着 γ 的增加, d 有所变大, 当 $\gamma=2.5$ 时, d 平均值稍大于 8. 这可以解释为, 随着 γ

的增加, 大度地标的平均度和它们之间的连接密度下降, 故直径也随之增大. 我们也在 Internet AS 图上进行了仿真. 对于 AS 图, d 仅为 2, 原因来自两个方面. 一方面是 Internet AS 图中存在少数度很大的节点, 另一方面是 Internet AS 图存在 Rich-Club 现象^[34], 即度较大的节点之间倾向于互相连接, 这些都会导致 d 比同等条件下的随机幂律图要小.

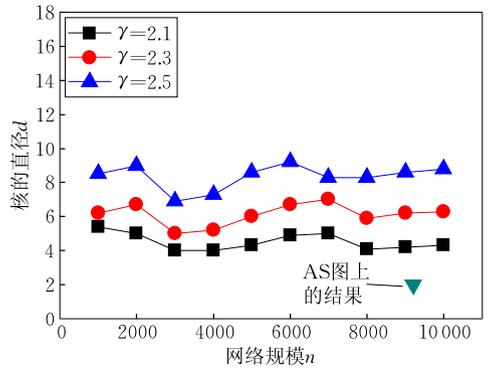


图 5 幂律图上 d 的测试结果

图 6 和图 7 比较了 NIHDLR 方法和 Abraham 方法在幂律图上产生的平均拉伸系数. 对于 NIHDLR, 我们分别测试了原始方法的平均拉伸系数和使用握手机制优化后的平均拉伸系数. 可以看出, NIHDLR 方法的平均拉伸系数显著低于 Abraham 方法. 特别是对于优化后的 NIHDLR 方法, 平均拉伸系数一般在 1.1 以下. 图 6 还给出了 Abraham 方法和 NIHDLR 方法在 AS 图上的测试结果. 前者的平均拉伸系数为 1.50 左右, 而后者在优化后只有 1.09. 从图 7 还可以观察到, 随着幂指数的增加, 原始 NIHDLR 方法的平均拉伸系数在增长, 而 Abraham 方法的平均拉伸系数有所降低. 这是因为, NIHDLR 方法是将节点的位置信息放在地标上, 当幂指数增大时, 度分布越均匀, 地标之间的连接越松散, 即地标之间的距离越大, 导致定位和路由过程需要付出更多的路径代价, 从而使整体的平均拉伸系数变大; 而 Abraham 方法将位置信息放在节点的邻域中, 这就比较适应度相对均匀的网络, 当幂指数增大时, 网络平均距离增大, 反而可以降低平均拉伸系数. 优化后的 NIHDLR 方法产生的平均拉伸系数则似乎随幂指数的增加没有多少变化, 原因是不需要在握手后再去查询目的地址, 所以较少受到地标之间距离增大的影响.

① The Cooperative Association for Internet Data Analysis (CAIDA). <http://www.caida.org/data>, 2013, 10, 23

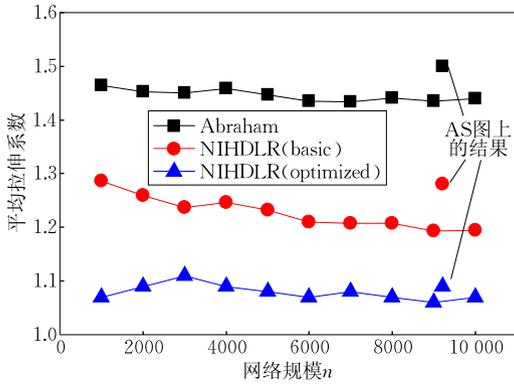


图 6 平均拉伸系数随网络规模变化的曲线

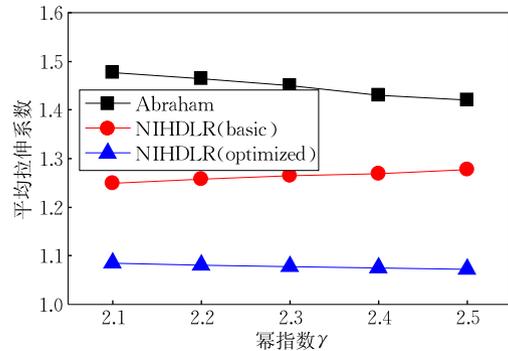


图 7 平均拉伸系数随幂指数变化的曲线

5.2 路由表大小

图 8 比较了 NIHDLR 方法和 Abraham 方法在幂律图上产生的平均路由表大小(注意:原始 NIHDLR 和优化后的 NIHDLR 在路由表设置方面并没有差别).为了简单起见,我们使用项数作为路由表大小的单位.原因是路由表项使用的比特数取决于实际的编码方法,不同的编码方法可能导致结果不一样.可以看出,NIHDLR 方法产生的平均路由表大小远小于 Abraham 方法产生的平均路由表大小.在 10 000 个节点的随机幂律图上,NIHDLR 方法的平均的路由表项数仅为 106 左右,而 Abraham 方法的平均的路由表项数接近 1500.在 AS 图上,NIHDLR 方法和 Abraham 方法的平均路由表项数分别为 101 和 1465 左右. NIHDLR 方法

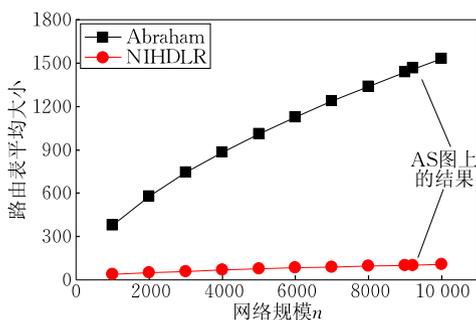


图 8 幂律图上的平均路由表大小

能产生很小路由表的原因是,在幂律图上对每个节点 $u \in V(G)$,由于 u 与大度节点有很高的概率直接相连或者只有很少的跳数, $B(u)$ 和 $C(u)$ 的规模都可以很小.而 Abraham 方法为了保证在一个节点的邻域内就能查询到任意目的节点的地址,往往导致较大的邻域.而节点的路由表必须维护到邻域节点的路由信息,因此路由表也较大.

6 讨论

目前我们的工作聚焦于为静态网络设计路由算法,即假设网络拓扑不会发生变化.然而实际网络的拓扑结构可能会发生变化,如 Internet 会存在链路或者节点失效的情况,失效的节点和链路也可以重新恢复.拓扑变化对本文提出的路由方法的影响和可能的应对策略讨论如下:

(1) 拓扑变化可能会引起某些节点地址的失效,因而需要对节点重新编址.因为节点的地址取决于离其最近的地标以及地标上与它连接的端口,因此拓扑变化可能也会引起这些部分的变化.一种方法是让所有地标定期地广播消息对网络进行宽度优先遍历,在消息中携带距离和端口号.每个节点在收到所有地标的消息后,选择最小距离的地标名及相应端口,对地址进行更新.

(2) 拓扑变化可能会引起某些节点路由表项的失效,因此需要利用变化后的拓扑进行更新.例如,节点到地标的最短路径和到其附近节点(即其 ball 和 cluster 中的节点)的最短路径均可能会因为拓扑变化而改变,因此对应的路由表项也应该更新.更新节点到地标的路由,可以通过上述的由地标发起的周期性网络遍历来完成.为了更新节点到其附近节点的路由,可以由该节点在其 ball 内发起小范围内的宽度优先遍历,同样在消息中携带距离和端口号,收到消息的节点在更新路由表的同时,可以对源节点发送一个应答,便于让源节点更新到它的路由表项.

(3) 地标失效的问题.由于地标充当了查询目的地址和路由转发的核心作用,因此地标失效也会引起路由失败.一种解决方法是在地标上存放冗余的节点名字/地址映射信息,一个节点可以映射到多个地标.当因为原来的地标失效导致路由失败时,源节点可以通过使用替代的地标来查询目的地址或者转发消息.

不可避免地,上述问题和措施会造成路由算法的复杂性增加和更多的性能开销.最近已经有工作考虑了如何在动态网络拓扑上改编紧凑路由算法以提高其适应性和实用性,并且通过实验证明改编的紧凑路由算法在拓扑变化时可以产生比传统路由算法(如链路状态路由算法)更低的消息开销^[35].然而,它是针对无线传感器网络的.无标度网络上的动态紧凑路由算法可能成为未来紧凑路由研究的一个热点.

7 结 论

本文针对类 Internet 拓扑结构的无标度网络提出了一种名字无关的紧凑路由方法.分析表明,该方法可以在无标度网络上取得比较优化的路由表大小和拉伸系数等性能指标.在无标度网络模型和真实 Internet AS 级拓扑图上的仿真实验表明,实际的拉伸系数和路由表大小均要显著优于最优的通用名字无关紧凑路由方法.我们注意到,最近有许多互联网专家认为,互联网路由系统应当采用位置符/标识符分离(Locator/ID Separation)的编址和路由方案^①.从本质来看,名字无关的紧凑路由方法可以看成是这类设想的一个实现.目前我们对 NIHDLR 的评估仅局限它在静态网络拓扑上的性能.在动态网络拓扑上,如何改编 NIHDLR 以提高它的动态适应性和实用性是我们下一步要考虑的工作.

致 谢 评审专家对本篇论文提出了宝贵意见,他们的建议对于提高本文的质量很有帮助.编辑老师付出了辛勤劳动,在此一并表示感谢!

参 考 文 献

- [1] Lin Chuang, Tian Yuan, Yao Min. Green network and green evaluation: Mechanism, modeling and evaluation. Chinese Journal of Computers, 2011, 34(4): 593-612(in Chinese)
(林闯,田源,姚敏.绿色网络和绿色评价:节能机制、模型和评价.计算机学报,2011,34(4):593-612)
- [2] Tang Ming-Dong, Zhang Guo-Qing, Yang Jing, Zhang Guo-Qiang. Scalable routing for the Internet. Journal of Software, 2010, 21(10): 2524-2541(in Chinese)
(唐明董,张国清,杨景,张国强.互联网可扩展路由.软件学报,2010,21(10):2524-2541)
- [3] Zhang Wei, Bi Jun, Wu Jian-Ping. Scalability of Internet inter-domain routing. Journal of Software, 2011, 22(1):

84-100(in Chinese)

(张威,毕军,吴建平.互联网域间路由可扩展性.软件学报,2011,22(1):84-100)

- [4] Krioukov D, Claffy K, Fall K, Brady A. On compact routing for the Internet. ACM SIGCOMM Computer Communication Review, 2007, 37(3): 43-52
- [5] Tang Ming-Dong, Liu Jian-Xun, Zhang Guo-Qing. A survey on compact routing. Journal of Frontiers of Computer Science and Technology, 2011, 5(3): 193-207(in Chinese)
(唐明董,刘建勋,张国清.紧凑路由研究.计算机科学与探索,2011,5(3):193-207)
- [6] Krioukov D, Fall K, Yang X. Compact routing on Internet-like graphs//Proceedings of the IEEE INFOCOM 2004. Hong Kong, China, 2004: 209-219
- [7] Brady A, Cowen L. Compact routing on power-law graphs with additive stretch//Proceedings of the 8th Workshop on Algorithm Engineering and Experiments. Miami, USA, 2006: 119-128
- [8] Tang Ming-Dong, Zhang Guo-Qing, Yang Jing, Zhang Guo-Qiang. A compact routing scheme for scale-free networks. Journal of Software, 2010, 21(7): 1732-1743(in Chinese)
(唐明董,张国清,杨景,张国强.针对无标度网络的紧凑路由方法.软件学报,2010,21(7):1732-1743)
- [9] Chen W, Sommer C, Teng S H, Wang Y. Compact routing in power-law graphs//Proceedings of the International Symposium on Distributed Computing (DISC'09). Elche, Spain, 2009: 379-391
- [10] Enachescu M, Wang M, Goel A. Reducing maximum stretch in compact routing//Proceedings of the IEEE INFOCOM 2008. Phoenix, USA, 2008: 977-985
- [11] Norros I. Powernet: Compact routing on Internet-like random networks//Proceedings of the Next Generation Internet Networks. Aveiro, Portugal, 2009: 1-8
- [12] Cowen L J. Compact routing with minimum stretch//Proceedings of the 10th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA). Baltimore, USA, 1999: 255-260
- [13] Thorup M, Zwick U. Compact routing schemes//Proceedings of the 13th ACM Symposium on Parallel Algorithms and Architecture (SPAA). Heraklion, Greece, 2001: 1-10
- [14] Arias M, Cowen L J, Laing K A, Rajaraman R, Taka O. Compact routing with name independence//Proceedings of the 15th Annual ACM Symposium on Parallel Algorithms and Architectures(SPAA). San Diego, USA, 2003: 184-192
- [15] Abraham I, Gavoille C, Malkhi D, et al. Compact name-independent routing with minimum stretch. ACM Transactions on Algorithms, 2008, 4(3): 37
- [16] Gavoille C, Gengler M. Space-efficiency for routing schemes of stretch factor three//Proceedings of the International Colloquium on Structural Information and Communication

① Locator/ID separation protocol (LISP). <http://tools.ietf.org/html/rfc6830>, 2013, 1

- Complexity (SIROCCO). Monte Verita, Ascona, Switzerland, 1997: 162-175
- [17] Laing K A. Name-independent compact routing in trees. *Information Processing Letters*, 2007, 103(2): 57-60
- [18] Korman A. Improved compact routing schemes for dynamic trees//*Proceedings of the 27th Annual ACM SIGACT-SIGOPS Symposium on Principles of Distributed Computing (PODC)*. Toronto, Canada, 2008: 185-194
- [19] Lu H I. Improved compact routing tables for planar networks via orderly spanning trees//*Proceedings of the 8th International Computing and Combinatorics Conference*. Singapore, 2002: 57-66
- [20] Faloutsos M, Faloutsos P, Faloutsos C. On power-law relationships of the Internet topology. *Computer Communications Review*, 1999, 29: 251-262
- [21] Barabasi A, Albert R. Emergence of scaling in random network. *Science*, 1999, 286(5439): 509
- [22] Albert R, Jeong H, Barabasi A L. Dimeter of the worldwide Web. *Nature*, 1999, 401: 130-131
- [23] Ripeanu M, Iarmnitchi A, Foster I. Mapping the Gnutella network: Properties of large-scale peer-to-peer systems and implications for system design. *IEEE Internet Computing*, 2002, 6(3): 50-57
- [24] Zhang G Q, Zhang G Q, Yang Q F, et al. Evolution of the Internet and its cores. *New Journal of Physics*, 2008, 10(6): 123027
- [25] Zhang Guo-Qing. Discovery of Internet topology knowledge and its applications. *Journal on Communications*, 2010, 31(10): 18-25(in Chinese)
(张国清. 互联网拓扑结构知识发现及其应用. *通信学报*, 2010, 31(10): 18-25)
- [26] Strowes S D, Perkins C. Harnessing Internet Topological stability in thorup-zwick compact routing//*Proceedings of the INFOCOM 2012*. Orlando, USA, 2012: 2551-2555
- [27] Chen W, Sommer C, Teng S-H, Wang Y. A compact routing scheme and approximate distance oracle for power-law graphs. *ACM Transactions on Algorithms*, 2012, 9(1): 1-26
- [28] Agarwal R, Godfrey P B, Har-Peled S. Approximate distance queries and compact routing in sparse graphs//*Proceedings of the INFOCOM 2011*. Shanghai, China, 2011: 1754-1762
- [29] Tang M D, Zhang G Q, Lin T, Liu J X. HLDDBR: A name-independent compact routing scheme for power-law networks. *Computer Communications*, 2013, 36(3): 351-359
- [30] Lu L. Probabilistic Methods in Massive Graphs and Internet Computing [Ph. D. dissertation]. University of California at San Diego, San Diego, USA, 2002
- [31] Chung F, Lu L. The average distances in random graphs with given expected degrees. *Internet Mathematics*, 2002, 99(7): 15879-15882
- [32] Chernoff H. A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations. *Annals of Mathematical Statistics*, 1952, 23(4): 493-507
- [33] Tarjan R E, Yao A C. Storing a sparse table. *Communications of the ACM*, 1979, 22(11): 606-611
- [34] Zhou S, Mondragon R J. The rich-club phenomenon in the Internet topology. *IEEE Communications Letters*, 2004, 8(3): 180-182
- [35] Mao Y, Wang F, Qiu L, Lam S, Smith J. S4: Small state and small stretch compact routing protocol for large static wireless networks. *IEEE/ACM Transactions on Networking*, 2010, 18(3): 761-774



TANG Ming-Dong, born in 1978, Ph. D., associate professor. His current research interests include network science and technology.

LIU Jian-Xun, born in 1970, Ph. D., professor. His current research interests include service computing and cloud computing.

ZHANG Guo-Qing, born in 1965, Ph. D., professor. His current research interests include computer networks and network science.

YANG Jing, born in 1952, Ph. D., professor. His current research interests include computer networks and future Internet architecture.

Background

In real-world communication networks such as Internet, shortest path routing algorithms are often used to perform routing tasks on them. However, the shortest path routing algorithms will likely cause scalability problems to routing tables imposed on nodes. That is, when the network size

increases, the routing table size will also grow rapidly, thereby significantly degrade the routing performance of the network. Compact routing has been considered as a good candidate to fundamentally address the scalability issue of the routing infrastructure. The basic idea of compact routing is

to achieve a good tradeoff between routing table size and path length. A compact routing algorithm guarantees that the size of the routing table on each node is $o(n)$, and in compensation, relaxes the shortest path requirement. The quality of a routing path used by a compact routing algorithm is quantitatively measured by *stretch*, which is the ratio between the length of the routing path and the shortest path.

This work focuses on compact routing on scale-free networks. Since the Internet has been revealed to be a scale-free network because of its power-law topological property, compact routing on this sort of networks has attracted increasing research interests. Several specialized compact routing algorithms have been proposed for scale-free networks, and they have been either theoretically or empirically proved to be more efficient than universal compact routing algorithms when applied to scale-free networks such as the Internet graph.

However, so far, nearly all the specialized routing algorithms for scale-free networks are name-dependent, which indicates that node names need to encode topologically sensitive information useful for routing. Moreover, the node names have to be renamed upon topology changes. In contrast,

name-independent algorithms do not encode topology-dependent information in the node names, hence are more flexible in coping with topology changes. Motivated by this, we proposed a name-independent compact routing algorithm that is specialized for scale-free networks. Based on the random power-law graph model and the Internet AS graph, we gave both analytical and experimental evaluations to the stretch and routing table size of the proposed routing algorithm. The analytical and experimental evaluations showed that it produces much better tradeoff between routing table size and stretch on scale-free networks than that of the universal optimal compact routing algorithms. Our proposed routing algorithm was essentially an implementation of the Locator/ID separation (LIS) idea, which has been strongly recommended by most recent proposals for the future Internet routing architecture. In this work, all the performance evaluation of our proposed routing algorithm was undertaken in static network environment. Understanding how it behaves in the dynamic environment is critical for developing a practically scalable routing protocol, which could be the future research direction.