

针对无标度网络的紧凑路由方法*

唐明董^{1,2,3+}, 张国清¹, 杨景¹, 张国强¹

¹(中国科学院 计算技术研究所,北京 100190)

²(中国科学院 研究生院,北京 100049)

³(湖南科技大学 知识处理与网络化制造湖南省普通高校重点实验室,湖南 湘潭 411201)

Compact Routing Scheme for Scale-Free Networks

TANG Ming-Dong^{1,2,3+}, ZHANG Guo-Qing¹, YANG Jing¹, ZHANG Guo-Qiang¹

¹(Institute of Computing Technology, The Chinese Academy of Sciences, Beijing 100190, China)

²(Graduate University, The Chinese Academy of Sciences, Beijing 100049, China)

³(Key Laboratory of Knowledge Processing and Networked Manufacturing, Hu'nan University of Science and Technology, Xiangtan 411201, China)

+ Corresponding author: E-mail: tangmingdong@ict.ac.cn

Tang MD, Zhang GQ, Yang J, Zhang GQ. Compact routing scheme for scale-free networks. Journal of Software, 2010,21(7):1732-1743. <http://www.jos.org.cn/1000-9825/3582.htm>

Abstract: Routing table size and routing path length are two critical measures for evaluating a routing algorithm, and there exists a tradeoff problem between them. Compact routing refers to the design of routing algorithms obtaining relatively optimal tradeoffs between the above two measures. So far, researchers have proposed quite a few universal compact routing schemes, which have high optimized bounds on routing table size and path length for arbitrary network topologies. However, as real-world networks usually have specific topologies, the universal schemes are possibly sub-optimal on them for not exploiting the topological properties. Recent work discovered many real-world networks had scale-free topologies. By exploiting the power-law and strong clustering features, a compact routing scheme with additive stretch for this class of networks is presented in this paper. By separating a network into a backbone tree and shortcuts, this scheme ensures between any source node and destination node in a network, the routing path length is at most an additive factor of b longer than the shortest path between them, and the local routing table at each node is upper bounded by $O(c \log^2 n)$ bits, wherein b and c are parameters related to the network topology. Experimental results show that b and c have small values in scale-free networks, and the proposed scheme can achieve better average-case performance than known schemes.

Key words: compact routing; scale-free network; network topology; simulation; stretch

摘要: 衡量一种路由算法优劣的两个重要指标是路由表的大小和路径的长度,但这两个方面通常是互相矛盾的.紧凑路由(compact routing)研究旨在设计路由算法在这两个指标上获得优化的平衡(tradeoff).目前,已有许多学者针对任意拓扑的网络提出了普适(universal)的紧凑路由方法(compact routing scheme).但是,真实的网络都具有特定

* Supported by the National Natural Science Foundation of China under Grant Nos.60673168, 90818004 (国家自然科学基金)

Received 2008-07-09; Revised 2008-10-16; Accepted 2009-01-15

的拓扑,普适的紧凑路由方法并没有利用真实网络呈现的特定拓扑特征,因而在这类网络上未必能取得最优的性能.最近的研究发现,许多真实网络都具有无标度特征和强聚集特征,利用这两类拓扑特征,提出了一种针对这类网络的紧凑路由方法.该路由方法将网络看成是由一个骨干树和一些捷径组成,在任意源节点和目的节点之间路由,使用路径的长度不超过它们的最短路径长度加上一个整数 b .路由表大小限制在 $O(c \log^2 n)$ 比特,其中, b 和 c 是由网络结构决定的参数.实验结果表明,在无标度网络上, b 和 c 可以同时取较小的值.与以往的紧凑路由方法相比,该方法在平均性能上表现更好.

关键词: 紧凑路由;无标度网络;网络拓扑;仿真;伸长系数

中图分类号: TP393 文献标识码: A

路由是通信网络研究中最基础的问题之一.节点路由表所需的存储空间和路径的长短,是衡量一种路由算法优劣的两个主要指标,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(shortest-path routing scheme)保证了任何两个节点之间都使用最短路径进行路由,但在每个节点上需要存放到所有其他节点的路由信息,路由表项数为 $O(n)$, n 代表网络规模,当 n 很大时,内存空间需求将成为系统性能的瓶颈.以 Internet 为例,如果不考虑策略因素,BGP(border gateway protocol)协议基本上采用最短路径路由.因此,随着网络规模的增加,核心路由器的路由表规模也随之迅速膨胀,严重制约了 Internet 的扩展性^[1].紧凑路由(compact routing)的提出正是为了解决上述问题,它的核心思想是,在路由表的大小和路径的长度上进行平衡,即设计的路由算法在每个节点上的路由表大小为 $o(n)$.但是,路由表项数的减少必将以增加路由的路径长度为代价.这种路由代价通常用伸长系数(stretch)来衡量,即,最坏情况下,实际使用路径相对于最短路径的拉伸程度.各类紧凑路由算法的研究都旨在针对给定网络寻求一种路由方法,能够在路由表的大小和伸长系数这两者之间达到较为优化的平衡关系.

对任何紧凑路由方法,它的伸长系数可以用 (α, β) 来表示,是指从任意节点 u 路由到 v 所使用的路径长度最大为 $\alpha d(u, v) + \beta$, 其中, $d(u, v)$ 为 u 和 v 之间的最短路径长度.如果一种路由方法的伸长系数为 $(\alpha, 0)$, 表示它具有倍数型(multiplicative)伸长系数,通常也简写为 α , 而 $(1, \beta)$ 表示增量型(additive)伸长系数.早期的紧凑路由研究仅限于特定拓扑的网络,如环、树、网格等^[2,3], 20 世纪 90 年代以来,许多学者针对一般性网络的紧凑路由开展了研究. Cowen^[4] 第一个提出了伸长系数为 $(3, 0)$ 的普适(universal)紧凑路由方法,本地路由表规模为 $O(n^{2/3} \log^{4/3} n)$ (下面称为 Cowen 方法). Thorup 和 Zwick^[5] 将 Cowen 方法中的路由表规模降低为 $O(n^{1/2} \log^{1/2} n)$, 这是迄今为止在伸长系数和路由表大小上取得最优平衡的普适路由方法(下面称为 TZ 方法). Gavoille 和 Gengler^[6] 证明了不存在伸长系数小于 $(3, 0)$ 且每个节点上路由表项数为 $o(n)$ 的普适路由方法.

普适的紧凑路由方法给出了在任意拓扑网络上最坏情况下的伸长系数,然而实际网络所表现出来的可能要比最坏情况下的理论值好很多.调研在真实网络上的紧凑路由机制,对网络的路由算法设计具有重要的指导意义.近年来的工作发现,许多真实的大规模网络都是无标度网络^[7], 节点度分布满足幂律的特征.其中, Internet AS 级的拓扑作为无标度网络的一个典型吸引了研究者们广泛的兴趣.研究发现,它还具有强聚集(clustering)^[8]、异配性(dissortative mixing)^[9]、富人俱乐部(rich-club)^[10] 以及自相似(self-similarity)^[11] 等特征. Krioukov 等人^[12] 提出将 TZ 方法用于 Internet 路由,通过分析和验证发现, TZ 方法在像 Internet 这类无标度图上的平均倍数型伸长系数远小于最坏情况下的理论值,约为 1.1 左右.但是, TZ 方法毕竟是通用的,设计时并没有专门针对无标度网络,由此产生的一个问题是:如果考虑无标度网络的拓扑特征,是不是可以设计出性能更好的路由方法.我们针对上述问题开展了研究.

本文针对无标度网络提出了一种具有增量型伸长系数的紧凑路由方法——BSROUT(backbone and shortcuts based routing),伸长系数等于 $(1, b)$, 保证将路由表大小限制在 $O(c \log^2 n)$ 比特,其中, b 和 c 是由网络结构决定的参数.与以往的工作相比,本文工作具有以下特点和贡献:

- 1) 从路由的问题出发,提出了一种针对无标度网络的分解方法.将网络看成由一个骨干和许多捷径组成,其中骨干是以最大度节点为根的最短路径生成树,捷径是网络中不在骨干树中的连接;
- 2) 发现无标度网络的上述骨干树的节点度分布也满足幂律,发现捷径的分布与网络的聚集特性十分相

关,捷径的分布特征决定了 b 和 c 可以在无标度网络取较小的值;

- 3) 发现在骨干树的基础上增加由少数捷径构造的生成树,由此计算的节点平均距离迅速接近于在原网络上的平均距离.利用这一特点,BSROUT 方法在无标度网络上使用较小的路由表就能够取得很低的平均伸长系数.在 AS(autonomous system)拓扑上的实验结果表明,BSROUT 方法比其他几种方法表现得更好.

本文第 1 节给出紧凑路由的相关概念,然后针对无标度网络提出一种网络骨干和捷径的定义.第 2 节首先介绍 BSROUT 方法的实现基础:针对树的混合紧凑路由算法和利用网络骨干和捷径构造生成树集合的算法,然后介绍 BSROUT 方法的设计,证明该方法具有 $(1,b)$ 的伸长系数、 $O(c \log^2 n)$ 比特的路由表和节点标记.第 3 节是实验部分,给出在几种典型无标度网络上骨干树和捷径所具有分布特征,以及使用 BSROUT 方法获得的参数值和实际性能.第 4 节介绍相关工作,并与已有方法进行比较.第 5 节对本文的工作进行讨论.最后是总结.

1 定义

1.1 紧凑路由相关概念

任何网络都可以抽象为一个图,有关网络、路由和紧凑路由的定义描述如下:

网络(network)的定义:给定任意图 $G=(V,E)$, V 为顶点集, E 为边集, $|V|=n$, $|E|=m$, 以及一个边的权重映射函数 $w:E \rightarrow R^+$, 记 $d(u,v)$ 是任意图 G 中任意两点的最短路径长度.本文工作仅考虑无权图.

路由方法(routing scheme)的定义:路由方法是定义在图 G 上的一种分布式算法,每个节点根据包首部和节点上路由表的信息做出决策,保证能够从任意节点发送一个包到任意的目标节点.一种路由方法通常由两部分组成:一是预处理过程;二是转发过程.预处理过程建立路由转发过程所必须的分布式数据结构,如路由表等.

路由方法的好坏主要由两个指标来衡量:一是预处理过程建立的分布式数据结构所占的存储空间的大小,通常指路由表的大小;另一个是路由的效率,即任意两点的路由是否是按照代价最小的路径,用伸长系数(stretch)来衡量.伸长系数可以用 (α, β) 表示,是指从任意节点 u 路由到 v 所使用的路径长度最大为 $\alpha d(u,v) + \beta$, 其中, $d(u,v)$ 为 u 和 v 之间的最短路径长度.其他指标还包括节点标记长度、包首部长度等,在动态拓扑上还需要考虑收敛代价(convergence cost).本文工作针对静态拓扑的情况,主要研究路由表大小与伸长系数的平衡.

紧凑路由(compact routing)的定义:如果一种路由方法在每个节点上的路由表项数为 $o(n)$, 伸长系数具有较小的上限,节点标记和包首部长度最多为 $\log n$ 的多项式,该路由方法可以称为紧凑路由.

名字依赖(name-dependent)的紧凑路由和名字独立(name-independent)的紧凑路由:名字依赖的紧凑路由又称为标记路由(labeling routing),是指路由方法需要对网络的节点重新分配名字(或标记),在名字中或多或少蕴含了网络的拓扑信息.名字独立的紧凑路由是指节点名字可以是任意的,不受路由方法的限制,路由算法仅根据包中指定的目标节点名字就能够将包递交给目标节点.名字独立的方法通常可以在名字依赖方法的基础上加上名字映射功能实现.本文的路由方法属于名字依赖的紧凑路由.

1.2 无标度网络、骨干和捷径

无标度(scale-free)网络广义上是指节点度分布满足 $P(k) \sim k^{-\gamma}$ 幂律分布的一类网络,这类网络通常具有较小的直径和平均距离,约为网络规模的对数形式 $O(\log n)$.许多无标度网络还具有强聚集特征,即存在大量的三角形连接.比如社会网络中,朋友的朋友很可能也是朋友.文献[13]中提出了一种网络骨干概念,定义为使边的介数(betweenness)之和最大的生成树,发现在无标度网络中这种树的节点度分布也满足幂律.但是,计算边的介数比较复杂,获取上述网络骨干算法的时间复杂度至少为 $O(mn)$.从网络的路由问题出发,考虑到无标度网络的特点,我们提出了不同的骨干树定义,生成这种树的时间复杂度为 $O(m)$.

骨干(backbone):设 $G=(V,E)$ 为无权无向连通图, $|V|=n$, $|E|=m$. h 为 G 的度最大节点,以 h 为源点的单源最短路径生成树,称为图 G 的骨干,用 T_0 表示.

捷径(shortcut):图 G 中不在 T_0 中的边称为捷径,捷径的集合用 S 表示.对任意一条捷径 $s \in S$, 设 $s=(u,v)$, 其中,

u, v 为 s 在 G 中的两个端点. 定义 s 的长度为 u 和 v 在 T_0 中的距离, 即 $length(s) = d_{T_0}(u, v)$.

引理 1. 设 S 是无权无向连通图 G 的捷径集合, 则 $|S|=m-n+1$, 且对任意捷径 $s \in S$, 有 $length(s) \geq 2$.

证明: 根据定义, S 为 G 中所有不在 T_0 中的边, 所以有 $|S|=m-n+1$. 而对于任意捷径 $s \in S$, 由于它的两个端点在 T_0 中不直接相连, 根据捷径长度的定义, 有 $length(s) \geq 2$. \square

2 BSROUT 的设计与分析

定理 1. 设 $G=(V, E)$ 为无权无向连通图, $|V|=n, |E|=m, T_0$ 为 G 的骨干, S 为 G 的捷径集合, b 为非负整数. 设 S 中长度大于等于 $b+2$ 的捷径数量为 c , 则可以设计一种路由方法, 使伸长系数为 $(1, b)$, 路由表大小为 $O(c \log^2 n)$ 比特, 节点标记长度也为 $O(c \log^2 n)$ 比特.

下面将详细介绍该路由方法的设计过程并给出必要的证明, 为了便于理解, 对使用的已知算法也作了简单介绍.

2.1 一种针对树的紧凑路由算法

BSROUT 方法建立在一种混合的树路由方法之上^[14], 该方法综合了文献[5]中的树路由方法(下面称为 TZ 树路由方法, 注意与 TZ 方法区别)和文献[15]中针对树的距离标记(distance labeling)方法(下面称为 Peleg 方法), 保证在树上总是使用最短路径路由, 路由表大小为 $O(\log^2 n)$ 比特, 包首部大小也为 $O(\log^2 n)$ 比特.

TZ 树路由方法对每个节点使用 $O(\log^2 n)$ 比特的标记和包首部、 $O(\log n)$ 比特的路由表. TZ 树路由方法预处理算法和路由决策算法两个部分, 下面作简单描述, 读者可参考文献[5]获得细节. 它的预处理过程是:

- 1) 给定任意树 T , 从 T 中任意选择一个节点 r 作为根节点对 T 进行深度优先遍历, 按访问的顺序从小到大使用 $\{1, 2, \dots, n\}$ 中的数字对节点进行标记;
- 2) 为每个非叶节点指定一个唯一的 heavy 子节点, 即为后裔节点数最多的子节点(相同情况下任取一个), 根节点 r 也看成是 heavy 节点, 其他节点称为 light 节点;
- 3) 在每个节点 v 上存放路由表 $label_T(v) = (f_v, h_v, e_h, e_0)$, f_v 表示 v 的子树中最大的数字标记, h_v 表示 v 的 heavy 子节点标记, e_h 表示通向 heavy 子节点的边, e_0 表示通向 v 的父节点的边;
- 4) 对每个节点 v 重新标记为 $label_T(v) = (v, L(v))$, 其中, $L(v)$ 表示从根节点 r 到 v 的路径上所有连接 light 节点的边和相应的父节点.

上述预处理的时间复杂度为 $O(m)$. 为了清楚起见, 后面用分配给节点 v 的数字代表 v , 用 $label_T(v)$ 表示节点 v 的标记. 图 1 展示了用 TZ 树路由方法对有 12 个节点的树进行预处理的情况, 带阴影的节点表示为 heavy 节点. 从根节点 1 到节点 12 的路径上, 由于 10 和 12 都是 light 节点, 所以节点 12 最后标记为 $label_T(12) = (12, L(12))$, 其中, $L(12) = ((1, edge(1, 10)), (10, edge(10, 12)))$, $edge(i, j)$ 表示节点 i 连接节点 j 的边. 路由决策算法描述如下: 设一个首部为 $(v, L(v))$ 的包从节点 u 路由至目标 v , 当前到达的节点为 w , 则在 w 上执行过程:

- 1) 如果 $w=v$, 到达目标节点, 终止; 否则, 转步骤 2);
- 2) 判断 $v \in [w, f_w]$ 是否成立, 如果不成立, 说明 v 不是 w 的后裔, 由边 e_0 转发包给 w 的父节点; 否则, 转步骤 3);
- 3) 判断 w 是否在 $L(v)$ 中出现, 如果是, 选择 $L(v)$ 中 w 对应的边转发包; 否则, 转步骤 4);
- 4) v 必是 h_w 或 h_w 的后裔节点, 由 e_h 转发给 h_w .

Peleg 方法对树采用递归分解的方法, 在 $O(n \log n)$ 时间内给全部节点分配长度为 $O(\log^2 n)$ 比特的标记, 根据标记可以在 $O(\log n)$ 的时间内准确计算出任意两个节点之间的距离. 下面用 Peleg 方法对 TZ 树路由方法进行改造, 从而得到一种在标记中包含距离信息的树路由算法——TZ' 算法.

设 T 为任意的无权树, 使用 TZ 树路由方法预处理后, 节点 v 的标记表示为 $label_T(v)$, v 的本地路由表为 $table_T(v)$. 用 Peleg 方法对 T 进行处理, 对每个节点 v , 设 $l'_T(v)$ 为 Peleg 方法分配给 v 的标记. TZ' 算法将 $l'_T(v)$ 与 $label_T(v)$ 合并得到 v 的新标记, 用 $label'_T(v)$ 表示, 同样将 $l'_T(v)$ 加入到 $table_T(v)$, 用 $table'_T(v)$ 表示. TZ' 方法在 T 中路由的路由决策算法与 TZ 树路由方法相同.

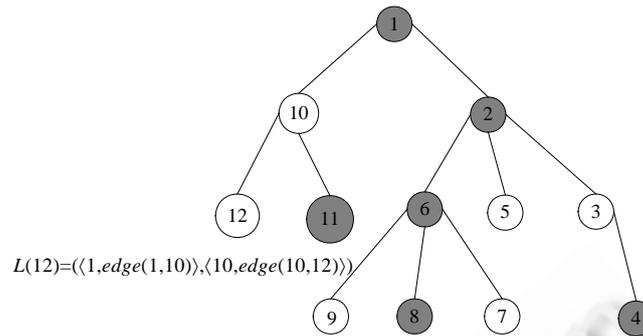


Fig.1 An example of the TZ tree routing scheme

图 1 TZ 树路由方法示例

引理 2. TZ'方法在每个节点上使用 $O(\log^2 n)$ 比特的标记和路由表以及 $O(\log^2 n)$ 比特的包首部.

证明:根据 TZ'方法的描述,很容易证明,故省略. \square

2.2 构造生成树集合的方法

将图的路由问题转化为树的路由是很多紧凑路由研究所采用的方法.BSROUT 方法需要从图 G 构造一种有效的生成树集合 TS .设 T_0 为 G 的骨干树, S 为捷径的集合,对任意捷径 $s \in S$,用 $length(s)$ 表示 s 的长度, $length(s)$ 等于 s 的两个端点在 T_0 上的距离.设 b 为给定非负整数,定义 $S' = \{s | length(s) \geq b+2, s \in S\}$ 表示长度大于或等于 $b+2$ 的捷径集合,定义 $c = |S'|$.下面是构造生成树集合 TS 的过程,时间复杂度为 $O(cm)$.

Procedure $TSConstruction()$.

将 T_0 加入到 TS 中;

$i \leftarrow 1$;

For each $(u,v) \in S'$

If TS 中不存在一棵生成树以 u 或 v 为根 then

以 u 为根生成一棵包括 (u,v) 在内的单源最短路径树,记为 T_i ;

将 T_i 加入到 TS 中;

$i \leftarrow i+1$;

End If

End For

引理 3. 设 TS 是根据上述算法构造的生成树集合,则 $|TS| \leq c+1$.

证明:已知 $|S'| = c$,从 S' 中的一条捷径最多构造一棵生成树,故 $|TS| \leq |S'| + 1 = c+1$. \square

2.3 图 G 上的紧凑路由方法

建立在上述部分的基础上,BSROUT 方法的设计分成 4 个部分:预处理部分、标记部分、存储部分和路由部分.

预处理部分包括图 G 的生成树集合 TS 的构造,并使用 TZ'方法对每棵树 $T_i \in TS$ 进行处理,对 T_i 中的每个节点 v ,赋予标记 $label'_i(v)$,路由表为 $table'_i(v)$,这部分为标记和存储部分提供了初始的数据结构.

标记部分对图 G 中每个节点 v 作如下处理:设 $label'_i(v)$ 是 v 在树 T_i 中的标记 ($0 \leq i < |TS|$),按 i 从小到大的顺序将 $(i, label'_i(v))$ 合成一个列表作为 v 的最终标记,记为 $label'_G(v)$.

存储部分对图 G 中每个节点 v 作如下处理:设 $table'_i(v)$ 是 v 在树 T_i 中的路由表 ($0 \leq i < |TS|$),按 i 从小到大的顺序将 $(i, table'_i(v))$ 合成一个列表作为 v 的最终路由表,记为 $table'_G(v)$.

路由部分的算法为:设从源节点 u 发送一个包到目标节点 v , 已知节点 v 的标记 $label'_G(v)$, 路由过程如下:

- 1) 在源节点 u 上, 对所有 $0 \leq i < |TS|$, 分别从 $table'_G(u)$ 和 $label'_G(v)$ 中抽取 $l'_T(u)$ 和 $l'_T(v)$, 计算 u 和 v 在 T_i 中的距离, 记为 $d_{T_i}(u, v)$;
- 2) 设 T_j 是所有 $0 \leq i < |TS|$ 中使 $d_{T_i}(u, v)$ 最小的树, 从 $label'_G(v)$ 抽取 $(j, label'_{T_j}(v))$, 然后使用 TZ' 算法在 T_j 中从 u 路由到包到 v .

引理 4. BSROUT 方法需要的预处理时间为 $O(cm+cn \log n)$.

证明: 预处理过程主要包括 TS 的构造和对 TS 中的每棵树用 TZ' 方法进行处理. 由第 2.2 节可知, 构造 TS 的时间为 $O(cm)$, 对每棵树用 TZ 树路由方法进行处理的时间为 $O(m)$, 用 Peleg 方法处理的时间为 $O(n \log n)$, 因此, 总的预处理时间为 $O(cm+cn \log n)$. \square

引理 5. BSROUT 方法对图中的每个节点使用 $O(c \log^2 n)$ 比特的路由表和标记.

证明: 对图 G 中的每个节点 v , 它在 TS 的全部生成树中出现. 因此, v 的标记 $label'_G(v)$ 包含 $|TS|$ 项 $(i, label'_{T_i}(v))$, 由引理 2 可知, 每项占 $O(\log^2 n)$ 比特, 由引理 3 可知 $|TS| \leq c+1$, 因此, $label'_G(v)$ 的长度为 $O(c \log^2 n)$ 比特. 同理, 可以得出 $table'_G(v)$ 的大小为 $O(c \log^2 n)$ 比特. \square

引理 6. BSROUT 方法具有增量型伸长系数 $(1, b)$, 使用 $O(\log^2 n)$ 比特的包首部.

证明: 由路由部分可知, 当源节点 u 向目标节点 v 发送一个包时, 是选择 u 和 v 距离最小的树 T_j 并使用 TZ' 算法路由, 所以包首部长度为 $O(\log^2 n)$. 下面证明 BSROUT 方法具有增量型伸长系数 $(1, b)$, 即对 G 中的任意两个节点 u, v , 设 $path'(u, v)$ 为使用本方法从 u 路由到 v 的路径, $SP(u, v)$ 表示 u 和 v 在 G 中的最短路径, 令 $d(u, v) = |SP(u, v)|$, 证明 $|path'(u, v)| \leq d(u, v) + b$. 分 3 种情况来考虑:

- 1) 如果 u 和 v 在 G 中的最短路径 $SP(u, v)$ 中不包含捷径, 即 u 到 v 的最短路径的所有边都在图 G 的骨干树 T_0 中, 选择在 T_0 中路由, 那么有 $|path'(u, v)| = d(u, v) \leq d(u, v) + b$, 成立;
- 2) 如果 $SP(u, v)$ 中包含长度大于或等于 $b+2$ 的一条捷径 s , 设 $s = (u', v')$, 由第 2.2 节的生成树集合构造算法知, 必存在一棵树 $T_i \in TS$, 以 u' 或 v' 为根 (设为 u'), 即有 $d_{T_i}(u, v) \leq d_{T_i}(u, u') + d_{T_i}(u', v) = d(u, u') + d(u', v)$. 而 u' 位于最短路径 $SP(u, v)$ 上, 所以 $d(u, v) = d(u, u') + d(u', v)$, 得出 $d_{T_i}(u, v) = d(u, v)$. 由于 $d_{T_i}(u, v)$ 已是最小, 故 $|path'(u, v)| = d_{T_i}(u, v) = d(u, v) \leq d(u, v) + b$, 成立.
- 3) 如果 $SP(u, v)$ 中含有 k 条捷径且长度都小于 $b+2$, 考察 u 和 v 在 T_0 中的距离, 即 $d_{T_0}(u, v)$. 已知 $|path'(u, v)| \leq d_{T_0}(u, v)$, 首先证明和 v 在 T_0 中有一个不同于 u 和 v 的最近共同祖先, 因为如果 u 是 v 在 T_0 中的祖先, 或 v 是 u 在 T_0 中的祖先, 那么必有 $d_{T_0}(u, v) = d(u, v)$, 这与假设 $SP(u, v)$ 中包含捷径矛盾. 令 w 是 u 和 v 在 T_0 中的最近共同祖先, 故 $d_{T_0}(u, v) = d_{T_0}(u, w) + d_{T_0}(w, v) = d(u, w) + d(w, v)$, 下面证明 $d_{T_0}(u, v) \leq d(u, v) + b$.

因为 u 和 v 在 G 上的距离 $d(u, v)$ 不等于 $d_{T_0}(u, v)$, 则 $SP(u, v)$ 中必存在一段路径, 与 u 和 v 在 T_0 上的最短路径围成一个环, 如图 2 所示 (w 是 u 和 v 在 T_0 的最近共同祖先, 实曲线表示 T_0 上两个端点之间的最短路径, 连接 x 和 y 的虚曲线表示 x 和 y 在 G 上的最短路径). 且该段路径的两个端点分别在 $SP(u, w)$ 和 $SP(w, v)$ 上, 设为 x 和 y . 设 $s_i = (x_i, y_i)$ 是 $SP(x, y)$ 上的任意一条捷径, $1 \leq i \leq k$, 令 $SP_{T_0}(x_i, y_i)$ 为 x_i 和 y_i 在 T_0 上的最短路径, 则 $SP(x, y)$ 上必存在一条捷径 $s_j = (x_j, y_j)$, 使 $SP_{T_0}(x_j, y_j)$ 包含 w 节点. 用反证法证明这一点, 假设对所有 $1 \leq j \leq k$, 不存在 $SP_{T_0}(x_j, y_j)$ 包含 w 节点, 那么将 $SP(x, y)$ 上的全部捷径 s_j 用 $SP_{T_0}(x_j, y_j)$ 替换将得到 x 与 y 在 T_0 中的一条新路径, 该路径不包含 w , 这与 T_0 是树相矛盾. 所以, $SP(x, y)$ 上必存在一条捷径 $s_j = (x_j, y_j)$ 使 $SP_{T_0}(x_j, y_j)$ 包含 w 节点, 故存在 $d_{T_0}(x_j, y_j) = d_{T_0}(x_j, w) + d_{T_0}(w, y_j)$. 已知 $d_{T_0}(u, v) = d(u, w) + d(w, v)$, 因为

$$d(u, w) \leq d(u, x_j) + d(x_j, w) \leq d(u, x_j) + d_{T_0}(x_j, w), d(w, v) \leq d(w, y_j) + d(y_j, v) \leq d_{T_0}(w, y_j) + d(y_j, v),$$

所以有 $d_{T_0}(u, v) \leq d(u, x_j) + d_{T_0}(x_j, w) + d_{T_0}(w, y_j) + d(y_j, v) = d(u, x_j) + d_{T_0}(x_j, y_j) + d(y_j, v)$. 又因为

$$d(u, v) = d(u, x_j) + d(y_j, v) + d(x_j, y_j) = d(u, x_j) + d(y_j, v) + 1,$$

即 $d(u, x_j) + d(y_j, v) = d(u, v) - 1$,

所以 $d_{T_0}(u, v) \leq d(u, v) - 1 + d_{T_0}(x_j, y_j)$. 根据捷径的定义, s_j 的长度为 $length(s_j) = d_{T_0}(x_j, y_j)$, 它小于或等于 $b+1$, 所

以 $d_{T_0}(u,v) \leq d(u,v) - 1 + \text{length}(s_j) \leq d(u,v) + b$. 最后得出 $|\text{path}'(u,v)| \leq d_{T_0}(u,v) \leq d(u,v) + b$. 得证. □

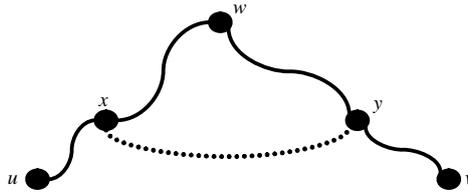


Fig.2 Illustration of the shortest paths between u and v in T_0 and G

图 2 u 与 v 分别在 T_0 和 G 上的最短路径图示

3 仿真实验

为了验证 BSRROUT 方法在无标度网络上的有效性和实际性能,我们分 3 个方面开展了仿真实验.首先测试了无标度图上骨干树的度分布和平均距离以及捷径长度的分布,证明该方法原则上是可行的.然后测试了该方法在无标度图上能取得什么样的 b 和 c ,即伸长系数和路由表规模的函数.最后测试了该方法在无标度图上的平均伸长系数.我们使用的无标度图为 BA(Barabási, Albert)模型、PFP(positive-feedback preference)模型和 Internet AS 级拓扑,其中,AS 拓扑数据是由 CAIDA^[16]组织提供的实际数据(ITDK0304),PFP 模型^[10]是目前建模 Internet AS 拓扑比较准确的模型,我们的实验恰恰也从另一方面对此作了验证.所有模型按平均节点度(k)=6 生成实例,每个测量值都是由 5 个实例计算后的平均值.

图 3 比较了无标度图和它们的骨干树的累积度分布,发现骨干树的节点度也呈现幂律分布,即具有无标度的特征,这意味着骨干树也具有较小的平均节点距离.通过计算得到骨干树平均距离和原图平均距离的比值,在 BA 图上约为 1.26,在 PFP 图上约为 1.21,在 AS 图上约为 1.17.这说明仅使用骨干树路由,就能得到比较低的平均伸长系数.

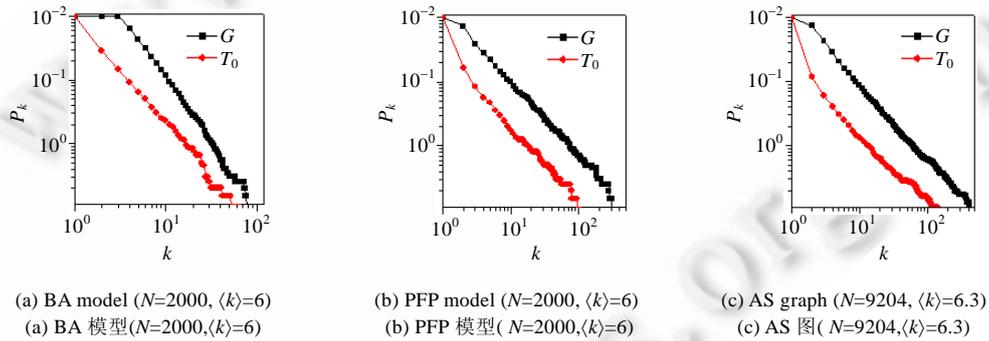


Fig.3 Cumulative degree distributions of scale-free graphs and their backbone trees

图 3 无标度图和骨干树的累积度分布

图 4 给出了无标度图上的捷径长度分布,横轴代表捷径长度,纵轴代表给定长度的捷径数量 m_s 占总捷径数量 m_0 的比例.对于 BA 模型和 PFP 模型,我们实验了节点数 $N=2000$ 和 $N=10000$ 两种情况.可以观察到,捷径长度分布在一个较小的范围内,且最大值较小.值得注意的是,BA 图与另外两种图的表现有所不同.对于 BA 图,捷径长度近似于呈泊松分布,大部分捷径位于中间区间,且这部分捷径的平均长度随着节点数 N 的增长有所增大.PFP 图和 AS 图的捷径长度分布十分相似,分布范围明显比 BA 图要小,捷径比例随捷径长度的增加锐减.特别地,随 N 的增长,捷径长度的分布范围没有明显变化,而且长度较大的捷径所占比例没有提高,反而有所下降.由于长度为 2 的捷径比例直接反映了图中三角形连接的规模,即反映了是否具有强聚集特征.因此从捷径长度的分布来看,BA 模型不具有强聚集特征,而 PFP 模型和 AS 图呈现强聚集特征,且随规模的增大而更加明显.捷

径长度的分布带给我们的启示是:伸长系数 b 可以在无标度图上取较小的值,且路由表不至于太大.具体来说,在 PFP 模型和 AS 图上, b 取 3 或 4 时,路由表项数约为长度大于等于 5 或 6 的捷径数目,这是可以接受的;而在 BA 图上, b 取可以接受的值与 N 有关,但 b 最多为 N 的对数形式,也是比较小的.

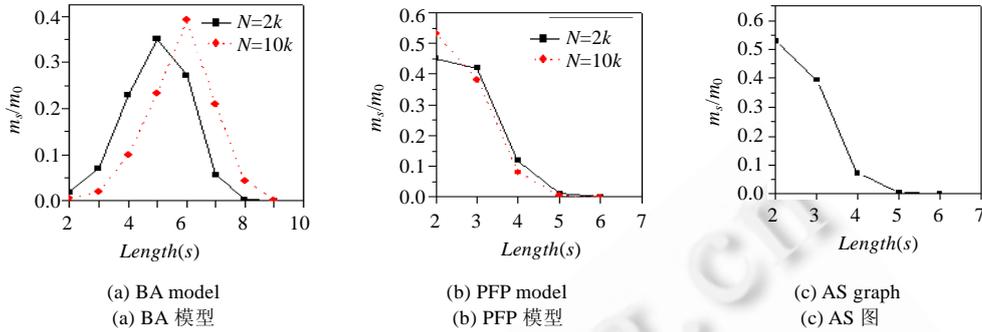


Fig.4 Length distributions of shortcuts for scale-free graphs

图 4 无标度图上捷径长度的分布

图 5 展示了本文方法中伸长系数等于 b 时,实际的路由表规模 $|TS|$ 相对于 b 的函数.前面阐述过, TS 是由长度大于 $b+1$ 的捷径构造而来的,且 $|TS| \leq c+1$.因此,图 5 与图 4 的实验结果是密切相关的. $b=0$ 时,表示所有路径都是最短路径.可以看出: b 超过某一阈值时, $|TS|$ 下降得很快;当 $b=4$ 时,AS 图中对应的 $|TS|$ 约为 2 左右.相对来说,BA 图上的结果要稍差一些,当 $b=6$ 时, $|TS|$ 约为 8 左右.这说明,在类似 Internet 的这种具有强聚集特征的无标度图上,伸长系数和路由表规模能够取得更优的平衡.

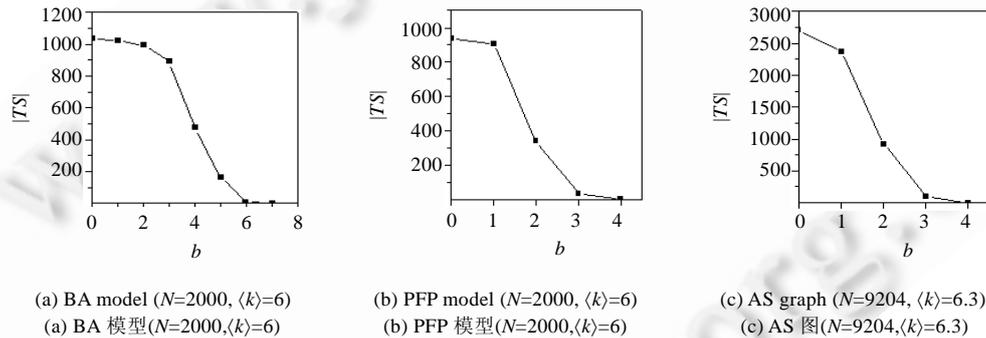


Fig.5 Relationship between $|TS|$ and b

图 5 伸长系数 b 与代表路由表规模的 $|TS|$ 之间的关系

图 6、图 7 给出了运行 BSRROUT 方法得到的平均伸长系数和使用的最短路径比例($N=2000, \langle k \rangle=6$).为了便于观察,将平均伸长系数用倍数型和增量型两种形式表示,前者表示使用路径的平均长度与最短路径平均长度的比值,后者表示使用路径的平均长度与最短路径平均长度的差值.具体步骤是:先计算仅使用 T_0 时的各种指标,然后根据捷径长度从大到小选择捷径增加少数的几棵树(不超过 20),每一步都测量所有节点对的使用路径长度和最短路径长度,然后求得上述指标.可以发现,使用很少的树就可以使平均伸长系数降至很低的值,且大部分使用路径是最短路径.对于 BA 图,当树的数目从 1 增加至 21 时,平均倍数型伸长系数从 1.26 降至 1.06,且近 80% 的路径使用最短路径;对于 PFP 图,当树的数目从 1 增加至 21 时,平均倍数型伸长系数从 1.21 降至 1.04,且近 90% 的路径使用最短路径.显然,使用同样多的树,PFP 图上的表现比 BA 图更好.这也说明,利用强聚集特征,BSRROUT 方法具有更高的实际性能.

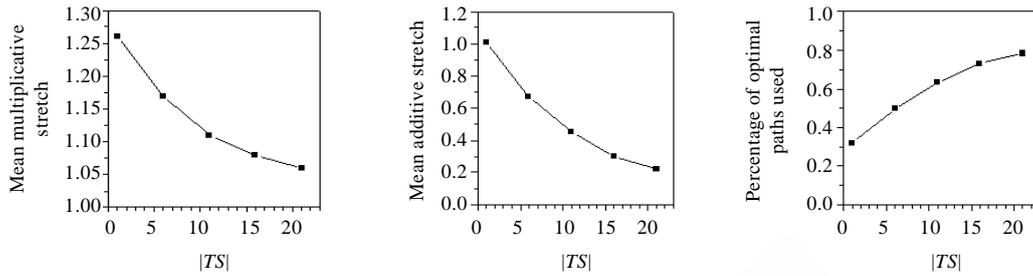


Fig.6 Observed average-case performance for BA graph

图 6 在 BA 图上运行得到的平均性能

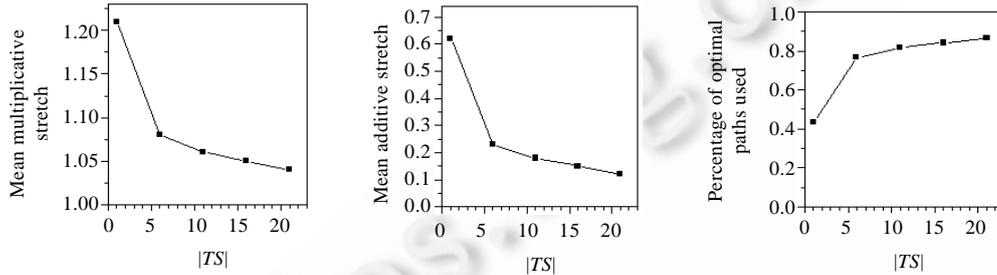


Fig.7 Observed average-case performance for PFP graph

图 7 在 PFP 图上运行得到的平均性能

4 相关工作与比较

对真实网络结构特征的研究近年来吸引了研究者们广泛的兴趣,其中对 Internet 结构特征的研究尤其是热点.文献[7]最早发现了 Internet AS 级拓扑存在无标度特征,此后,AS 拓扑的强聚集、异配性、富人俱乐部以及自相似等特征被其他学者陆续发现.文献[17]进一步发现,AS 拓扑在演化过程中呈现核心稳定、最大度稳定和边数弱加速增长等特征.文献[18]指出,AS 拓扑的一些特征之间存在较强的关联性.随着这些特征的发现,它们对网络功能的影响也引起了较多的关注.这里仅对无标度网络上利用拓扑特征改进路由的相关工作进行介绍,如文献[14,19,20]等,并给出了比较.

文献[19]提出的路由方法(下面称为 CRAMI 方法)与 TZ 方法有相似之处,但是利用了无标度网络的小世界特征,在节点上只存放邻居和到少数 hub 节点的路由信息,而在节点标记中包含更多的信息用于辅助路由.在无标度图上的实验结果表明,使用同样的平均路由表规模,CRAMI 方法的平均倍数型伸长系数比 TZ 方法低 0.023 左右.但是,CRAMI 方法没有给出最坏情况下伸长系数和路由表大小的界限.实际上,由于幂律的存在,少数高度的节点可能需要 $O(n)$ 的路由表.文献[14]提出了一种伸长系数为 $(1,d)$ 的紧凑路由方法(下面称为 BC 方法),路由表大小最多为 $O(e \log^2 n)$ 比特,它将网络分为内核和外围两部分,内核的半径为 $d/2$, e 是使外围成为森林所需要删除的边数.利用无标度网络外围连接比较稀疏的特点, d 和 e 都可以取较小的值.实验结果表明,平均伸长系数比 TZ 方法降低了 0.03 以上.但是,BC 方法中代表伸长系数的 d 只能取偶数.文献[20]则针对降低无标度网络上的最大伸长系数进行了研究,通过在无标度图上的实验,其结果表明,存在路由方法具有低于 $(3,0)$ 的伸长系数,并且在很大概率上使得路由表规模为 $o(n)$,但是并没有给出具体的路由方法设计.

本文的方法与 BC 方法属于一类方法,需要构造多个生成树,然后选择使源节点和目标节点距离最小的树进行路由,可以获得增量型的伸长系数.由于在无标度网络的树覆盖中节点距离与在网络中的距离接近,因此这类方法效果较好.与 BC 方法相比,我们采取了对无标度网络的不同分解方式,使用了不同的生成树集合的构造方法.我们将网络看成由一个骨干树和许多捷径组成,无标度网络上长度较大的捷径较少,因此在 b 和 c 上都能

够取得较小的值,特别是在类似 Internet 的这种具有强聚集特征的图上, b 的取值更小.我们的方法克服了 BC 方法中伸长系数只能取偶数的缺点,同时利用了网络的强聚集特征.

基于相同的 AS 拓扑数据(ITDK0304),将 BSROUT 方法的结果与文献[1]中给出的已知方法的运行结果进行了比较,平均伸长系数和路由表项数见表 1($N=9204, M=28959$).可见,BSROUT 方法在 AS 拓扑图上的表现要比其他几种方法更好,在路由表项数相同的情况下,平均伸长系数低于其他方法.我们认为可作如下解释:文献[21]的工作发现 AS 图中聚集系数 $C(k)$ 随节点度 k 的增加而递减,因此在 AS 图外围的低度节点上也存在很多三角形连接,BC 方法倾向于选择这些三角形中的某一条边创建生成树;而从我们的方法来看,这些边通常是作为长度为 2 的捷径,与选择这些长度较小的捷径相反,我们倾向于选择长度较大的捷径来构建生成树,更大程度上缩短了节点的平均距离,从而可以较快地降低路由的平均伸长系数.

Table 1 Comparison of performance between our scheme and other schemes on AS graph

表 1 本文路由方法与其他紧凑路由方法在 AS 拓扑图上的性能比较

Scheme	Mean multiplicative stretch and routing table size (in entries)
TZ ^[5]	1.08 stretch, 62 entries
BC ^[14]	1.06 stretch, 22 entries
TZ/BC hybrid ^[14]	1.02 stretch, 84 entries
BSROUT	1.03 stretch, 22 entries or 1.01 stretch, 50 entries

5 讨论

前面几节给出了本文算法的详细设计、分析和验证,这里将对本文工作获得的结果和在 Internet 环境下应用紧凑路由面临的问题进行讨论.

紧凑路由研究的动机之一是为像 Internet 这样的大规模网络设计可扩展的路由算法.Internet 的 BGP 路由表规模近年来一直呈现指数级增长,庞大的路由表导致了路由器开销的增大和网络整体性能的下降,严重制约了 Internet 的扩展性.紧凑路由可以大幅降低路由表的规模.但是,路由表规模的降低会导致伸长系数的增大.因此,衡量路由可扩展性的关键是路由算法能否同时使用尽可能小的路由表和伸长系数.本文提出的 BSROUT 方法在类 Internet 的无标度网络上平均伸长系数接近于 1,而路由表规模是 $O(c)$ 项或 $O(c \log^2 N)$ 比特, c 能够取较小的值,如 $\log N$ 的多项式形式.也就是说,在像 Internet 的这类网络上,即使每个 IPv4 地址都作为一个节点,即 $N=2^{32}$,该方法在每个节点上也只需用规模为 $10^2 \sim 10^3$ 数量级的路由表.而目前 Internet 上核心 BGP 路由器的路由表规模早已在 25 万项以上^[22].因此,BSROUT 方法在伸长系数上比 BGP 协议使用的路由算法平均只有很小的增长,但是在路由表规模上都可以支持网络的“无限”扩展.

针对 Internet 设计可扩展的路由算法,优化路由表规模和伸长系数是必要的第一步.在真实的 Internet 环境下设计路由算法,还需要考虑以下因素,如节点名字与拓扑独立、拓扑动态变化、AS 关系以及策略等.迄今为止,紧凑路由研究只考虑静态拓扑的情况,以名字依赖的算法为主,侧重于提高路由的静态性能.因此,上述因素对紧凑路由的影响还需要进一步探讨.节点名字与拓扑独立,即名字独立的紧凑路由算法可以通过在名字依赖的算法基础上增加名字映射功能得到,这可能会使路由表规模和伸长系数有一定的增长.值得注意的是,文献[23]提出了一种普适的名字独立的紧凑路由方法,它与 TZ 方法的伸长系数和路由表规模在理论上几乎保持了一致,这为我们提供了启发和思路.在本文工作的基础上,利用 Internet 拓扑特性设计优化的名字独立的紧凑路由方法是我们目前正在进行的工作.考虑到网络节点或链路可能出现故障,真实网络上的路由协议通常需要适应动态的拓扑.由拓扑变化引起的收敛代价也是衡量路由算法优劣的一个重要指标.BGP 协议在收敛代价上也存在扩展性问题,比如路由更新消息数量也一直呈现超线性增长、收敛慢等^[1].因此,针对 Internet 设计可扩展的路由算法还必须大幅降低收敛代价,使用较小的路由表理论上有利于降低路由更新总数并加速收敛.但是,针对这方面的紧凑路由研究还很少见.最后,Internet 路由需要满足 AS 关系和策略,紧凑路由如何支持这些功能是一个新的研究问题,目前还未见到相关的研究.

限于紧凑路由研究的现状,本文的工作虽然针对静态拓扑,但是对具体网络上的路由算法设计仍具有一定的

指导意义.一般认为,“结构决定功能”,路由作为网络的基本功能之一,利用网络结构特征必定能够优化路由算法的性能.本文的工作利用了网络的无标度特征以及强聚集这样更深层次的特征,并验证了这一点.

6 总结

本文针对无标度网络提出了一种紧凑路由方法,伸长系数为 $(1,b)$,路由表大小为 $O(c \log^2 n)$ 比特,其中 b 和 c 是由网络拓扑决定的参数.该方法利用了无标度网络的幂律和强聚集特征.实验结果表明,在 b 和 c 上可以同时取较小的值.与以往的紧凑路由方法相比,该方法在路由表大小和伸长系数的平衡上表现更好.对于许多真实的大规模网络,如 Internet,过于庞大的路由表常常成为系统性能的瓶颈.因此,紧凑路由可以提高网络的可扩展性.

目前的紧凑路由研究以理论为主,只考虑静态网络.下一步将在两个方面开展研究:一方面,真实网络除了无标度和强聚集的拓扑特征,还具有许多其他拓扑特征,这些拓扑特征对路由性能的影响值得进一步研究;另一方面,从应用出发研究紧凑路由在 Internet 环境下的可行性,如设计适应动态拓扑、AS 关系以及策略等因素的紧凑路由方法.

References:

- [1] Krioukov D, Claffy KC, Fall K, Brady A. On compact routing for the Internet. ACM SIGCOMM Computer Communication Review, 2007,37(7):43–52.
- [2] Santoro N, Khatib R. Labeling and implicit routing in networks. The Computer Journal, 1985,28(1):5–8. [doi: 10.1093/comjnl/28.1.5]
- [3] Leeuwen J, Tan RB. Interval routing. The Computer Journal, 1987,30(4):298–307. [doi: 10.1093/comjnl/30.4.298]
- [4] Cowen L. Compact routing with minimum stretch. Journal of Algorithms, 2001,38(1):170–183. [doi: 10.1006/jagm.2000.1134]
- [5] Thorup M, Zwick U. Compact routing schemes. In: Proc. of the 13th ACM Symp. on Parallel Algorithms and Architecture (SPAA 2001). Heraklion: ACM Press, 2001. 1–10.
- [6] Gavoille C, Gengler M. Space-Efficiency for routing schemes of stretch factor three. Journal of Parallel and Distributed Computing, 2001,61(5):679–687. [doi: 10.1006/jpdc.2000.1705]
- [7] Faloutsos M, Faloutsos P, Faloutsos C. On power-law relationships of the Internet topology. ACM SIGCOMM Computer Communications Review, 1999,29(4):251–262. [doi: 10.1145/316194.316229]
- [8] Dorogovtsev SN. Clustering of correlated networks. Physical Review E, 2004,69(027104). [doi: 10.1103/PhysRevE.69.027104]
- [9] Newman MEJ. Assortative mixing in networks. Physical Review Letter, 2002,89(208701). [doi: 10.1103/PhysRevLett.89.208701]
- [10] Zhou S, Mondragon RJ. Accurately modeling the Internet topology. Physical Review E, 2004,70(066108). [doi: 10.1103/PhysRevE.70.066108]
- [11] Zhou S, Zhang GQ, Zhang GQ. Chinese Internet AS-level topology. IET Communications, 2007,1(2):209–214. [doi: 10.1049/iet-com:20060518]
- [12] Krioukov D, Fall K, Yang X. Compact routing on Internet-like graphs. In: Proc. of the IEEE INFOCOM 2004. Hong Kong: IEEE, 2004. 209–219.
- [13] Kim DH, Noh J D, Jeong H. Scale-Free trees: The skeletons of complex networks. Physical Review E, 2004,70(046126).
- [14] Brady A, Cowen L. Compact routing on power-law graphs with additive stretch. In: Proc. of the 8th Workshop on Algorithm Engineering and Experiments. SIAM, 2006. 119–128. <http://www.siam.org/meetings/alnex06/>
- [15] Peleg D. Proximity-Preserving labeling schemes and their applications. In: Proc. of the 25th Int'l Workshop on Graph-Theoretic Concepts in Computer Science. LNCS 1665, Springer-Verlag, 1999. 30–41.
- [16] CAIDA website. <http://www.caida.org>
- [17] Zhang GQ, Zhang GQ, Yang QF, Cheng SQ, Zhou T. Evolution of the Internet and its cores. New Journal of Physics, 2008, 10(123027). [doi: 10.1088/1367-2630/10/12/123027]
- [18] Zhang GQ, Zhang GQ. Research on Internet correlation. Journal of Software, 2006,17(3):490–497 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/17/490.htm> [doi: 10.1360/jos170490]

- [19] Carmi S, Cohen R, Dolev D. Searching complex networks efficiently with minimal information. *Europhysics Letters*, 2006,74(6): 1102–1108. [doi: 10.1209/epl/i2006-10049-1]
- [20] Enachescu M, Wang M, Goel A. Reducing maximum stretch in compact routing. In: *Proc. of the IEEE INFOCOM 2008*. Phoenix: IEEE, 2008. 977–985.
- [21] Zhang GQ, Zhang GQ. Exploring the local connectivity preference of the Internet AS-level topology. In: *Proc. of the IEEE Int'l Conf. on Communications (ICC 2007)*. Glasgow: IEEE, 2007. 6439–6445.
- [22] BGP routing table analysis reports. 2008. <http://bgp.potaroo.net/>
- [23] Abraham I, Gavoiile C, Malkhi D, Nisan N, Thorup M. Compact name-independent routing with minimum stretch. *ACM Trans. on Algorithms*, 2008,4(3):37.

附中文参考文献:

- [18] 张国强,张国清. Internet 网络的关联性研究. *软件学报*, 2006,17(3):490–497. <http://www.jos.org.cn/1000-9825/17/490.htm> [doi: 10.1360/jos170490]



唐明董(1978—),男,湖南祁阳人,博士,讲师,主要研究领域为计算机网络,图论应用.



杨景(1952—),男,博士,研究员,博士生导师,主要研究领域为网络业务控制,集成体系结构.



张国清(1965—),男,博士,CCF 高级会员,主要研究领域为计算机网络,网络结构.



张国强(1980—),男,博士,主要研究领域为计算机网络.