

基于地标节点覆盖面的TZ紧凑路由算法研究

秦晓伟,周明洋,嵇钊,傅忠谦

(中国科学技术大学电子科学与技术系,安徽合肥 230027)

摘要:通过地标节点选取机制,TZ紧凑路由算法很好地保证了路由系统的扩展性,但TZ紧凑路由算法并没有限制地标节点的覆盖面,也没分析覆盖面过小的地标节点是否利于信息的传递.本文研究发现覆盖面过小的地标节点不利于紧凑路由的性能,因此通过限制地标节点的覆盖面,并在地标节点选取过程中删除覆盖面过小的地标节点,改进了TZ紧凑路由算法;同时,系统地分析了地标节点的覆盖面与平均伸长系数、平均路由表的关系.在连续10年的Internet AS图上进行仿真,实验结果表明,随着地标节点最小覆盖面的增大,平均伸长系数先减小而后逐渐增加,平均路由表先减小而后保持不变;当选取一个合适的阈值时,改进的算法比原始算法有更小的平均伸长系数和平均路由表,有效提升了紧凑路由的性能.

关键词:紧凑路由; TZ算法; Internet AS图; 地标节点; 覆盖面

中图分类号: TP301 **文献标识码:** A doi:10.3969/j.issn.0253-2778.2015.07.010

引用格式: QIN Xiaowei, ZHOU Mingyang, ZHUO Zhao, et al. Study on TZ compact routing schemes based on the coverage of landmark[J]. Journal of University of Science and Technology of China, 2015, 45(7): 594-600.
秦晓伟,周明洋,嵇钊,等. 基于地标节点覆盖面的TZ紧凑路由算法研究[J]. 中国科学技术大学学报, 2015, 45(7): 594-600.

Study on TZ compact routing schemes based on the coverage of landmark

QIN Xiaowei, ZHOU Mingyang, ZHUO Zhao, FU Zhongqian

(Department of Electronic Science and Technology, University of Science and Technology of China, Hefei 230027, China)

Abstract: TZ compact routing algorithm guarantees the scalability of routing system by the mechanism of selecting landmarks, but though it does not limit the coverage of landmark or analyse whether the landmarks with small coverage are beneficial for the message transfer. It was found that the landmarks with small coverage go against the performance of compact routing, so landmarks were deleted whose coverage was less than the threshold to keep the coverage of landmarks. Also, its relationships with the average stretch and the average routing table size were analyzed systematically. Applying TZ compact routing algorithm to the snapshots of the AS graph spanning a 10 year period, it was found the average stretch decreases first and then increases gradually, the average routing table size decreases first and then keeps invariant. When choosing an approximate threshold, the improved algorithm shows lower average stretch, lower average routing table size and which effectively improves the performance of TZ compact routing.

Key words: compact routing; TZ algorithm; Internet as graph; landmark; coverage

收稿日期:2014-12-16;修回日期:2015-03-06

作者简介:秦晓伟,男,1990年生,硕士生.研究方向:复杂系统和网络理论. E-mail:syzf@mail.ustc.edu.cn

通讯作者:傅忠谦,博士/副教授. E-mail:zqfu@ustc.edu.cn

0 引言

当前,互联网路由系统的扩展性正面临严重的挑战.对于传统的 BGP 最短路径路由协议,Internet AS 层上每个节点要维护到所有节点的信息,且 AS 层上节点数量从 2002 年 10 000 个到 2010 年 34 000 个,并且仍在呈线性增长^[1].庞大的路由表项将显著增加路由表的内存与处理器开销,导致通信时延恶化等一些列问题.而这将引起互联网路由系统的扩展问题^[2-3].紧凑路由算法^[4]抛弃了传统的最短路径传输要求,通过将一部分节点作为信息中转节点(地标节点),在路由长度增加不大的情况下大幅降低路由表长度,从而在路由表大小和路径长度间取得平衡,保证了路由系统的扩展性.

对于紧凑路由的研究,目前学术界旨在寻找一种路由算法,来尽可能优化路由表和路径长度间的平衡关系. Cowen 首次提出伸长系数为 3,本地路由表为 $O(n^{2/3})$ 的路由算法(伸长系数为网络数据包实际的传输路径与最短路径的比值)^[5]. Thorup 等改进了 Cowen 的算法,提出了伸长系数为 3,本地路由表为 $O(\sqrt{n/\log n})$ 的路由算法(TZ 算法)^[6].利用 Internet AS 图的幂律特性^[7], Brady 等提出了一种具有增量型伸长系数(1, d)的紧凑路由算法(BC 算法)^[8].对于合成的幂率随机图和 Internet AS 图, TZ 算法和 BC 算法都呈现非常好的性能,其平均伸长系数只有 1.1 左右,远小于其理论边界^[9-10].同样,中国科学院计算技术研究院的张国清等针对无标度网络提出了一种基于骨干树的紧凑路由算法^[11]. Stephen 等更改初始地标节点的选取机制,优化了 TZ 算法^[12-13].

TZ 算法是比较常用的紧凑路由算法,通过迭代过程高效地选取地标节点.由于 TZ 算法在迭代的过程中并没有限定地标节点的覆盖面,因此选取的地标节点的覆盖面会有差异性,甚至会产生部分覆盖面非常小的地标节点,这些覆盖面比较小的地标节点占据着路由表的大小却不一定利于全局数据包的转发.为此,本文系统地研究了地标节点的覆盖面与紧凑路由的关系.具体地说,通过设定一个阈值 k 来限制地标节点的覆盖面,并在迭代过程中剔除这些邻域小于 k 的地标节点;同时详细地分析了地标

节点的覆盖面与路由表、伸长系数的关系.在真实的 Internet AS 层数据上仿真发现,随着地标节点最小覆盖面 k 的递增,平均伸长系数先减小而后逐渐增大,平均路由表先减小而后保持不变,因此地标节点的覆盖范围与紧凑路由效率密切相关,通过选择合适的地标节点覆盖范围,可有效地提高紧凑路由算法的性能.

1 TZ 紧凑路由算法及其改进

1.1 TZ 算法和地标节点选取

对于实际网络,常用图模型来抽象网络的拓扑结构.给定一个无向连通图 $G(V, E)$, V 是节点的集合, E 是边的集合, $n = |V|$ 表示图中节点的个数.定义 $d(v, w)$ 为图 G 上节点 v 和节点 w 之间的最短距离(一般为跳数), L 为地标节点集合, $L(v) = \{w \in L \mid \min\{d(v, w)\}\}$ 表示节点 v 的地标节点,它是距离节点 v 最近的地标节点. $C(v)$ 表示节点 v 的邻域,它表示这样的节点集合:节点到节点 v 的距离小于节点到自身地标节点的距离.

$$C(v) = \{w \in V \mid d(v, w) < d(w, L(w))\} \quad (1)$$

TZ 算法选取地标节点是个迭代过程^[6],其地标节点的选取过程如下:首先,以概率 $p = s/|W|$ 随机地从集合 W 中选取 s 个节点并入到地标节点集合(初始时,地标节点集合为空,且 $W = V$).其中, $s = \sqrt{n/\log n}$;然后,用式(1)来构建每一个节点的本地邻域,对于邻域规模比较大的节点,根据式(2)选取本地邻域超过阈值 $4n/s$ 的节点集合来更新 W ;重复算法直到集合 W 为空.

$$W = \{v \in V \mid C(v) > 4n/s\} \quad (2)$$

算法迭代的最大次数为 $2\log n$ ^[6],故地标节点的规模为 $O(\sqrt{n/\log n})$.当算法收敛时,由式(2)可知,节点的邻域规模为 $O(\sqrt{n/\log n})$.节点路由表包括本地邻域节点集合和全局地标节点集合,故路由表的规模为 $O(\sqrt{n/\log n})$.

TZ 算法是通用的命名依赖的路由算法,因此需要对节点标记,且节点的标记与拓扑有关.节点标记包括三部分:目的节点标记、目的节点的地标节点标记、目的地标节点到目的节点的下一跳标记.路由包信息的时候,若目的节点不在本地路由表内,则路由包信息到目的节点的地标节点,该地标节点再用

下一跳标记将包信息传递给目的节点,反之,直接传递给目的节点.

1.2 地标节点特性分析

对紧凑路由而言,关键是选取合适的地标节点,而地标节点特征的最直接表现就是其覆盖面.对任意地标节点 $v \in L$,其覆盖面 $C_L(v)$ 定义为:

$$C_L(v) = \{w \in V \mid L(w) = v\} \quad (3)$$

对于 TZ 算法,在迭代过程中,节点的地标节点会不断的改变;同时初始地标节点是随机地选取的,且在迭代过程中并没有限制地标节点的覆盖面,因此,选择的地标节点的覆盖面会有差异性,甚至会有部分覆盖面比较小的地标节点,这些地标节点占据着路由表却不一定利于全局数据包的转发.

图 1(a)展示了 TZ 算法中部分地标节点的分布,其中黑色的节点表示地标节点,显然,地标节点的覆盖面分布不均衡,相比其他地标节点,有些地标节点的覆盖面很小,比如节点 4;图 1(b)显示了修改后的地标节点分布,其中覆盖面较小的地标节点(节点 4)被剔除了.

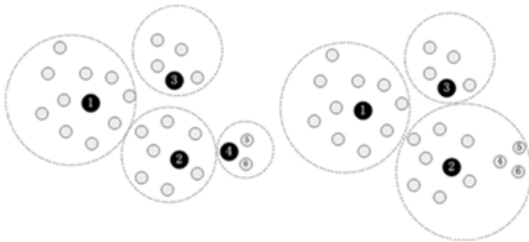


图 1 地标节点特性分析

Fig. 1 The property of landmarks

对于一个大规模的网络,期望整个网络的地标节点都能有图 1(b)这样的分布特性,同时保持较小的路由表和较小的伸长系数,因此很有必要研究地标节点的覆盖面与伸长系数、路由表大小的关系,并期望通过限制地标节点的覆盖面来提高紧凑路由的性能.

1.3 改进的 TZ 算法

对于 Internet,节点的度服从幂律分布^[7],度大的节点与其他节点有更好的连通性,因而更容易成为地标节点^[9],但这依然无法保证地标节点的覆盖面,因此需要设置一个阈值 k 来限制地标节点的覆盖面.此处,将添加阈值 k 来约束地标节点覆盖面的改进算法称为 TZ_k ,具体描述如下.

算法 1.1 Center(G, s, k)

```

L ← ∅, D ← ∅, W ← V
while(W ≠ ∅ && D ≠ W) do
{
  L ← L ∪ Sample(W, s)
  L(v) = {w ∈ L | min{d(v, w)}}, for every v ∈ V
  C_L(v) = {w ∈ V | L(w) = v}, for every v ∈ L
  L ← {v ∈ L | C_L(v) ≥ k}
  D ← {v ∈ L | C_L(v) < k}
  L(v) = {w ∈ L | min{d(v, w)}}, for every v ∈ V
  C(w) = {v ∈ V | d(w, v) < d(v, L(v))}
  W ← {w ∈ V | |C(w)| > 4n/s}
  if( W ≠ ∅ && |L| < s)
    W ← RandomOneMax(k)
}
Return L

```

算法中, s 表示初始时要选取的地标节点数目, k 表示限制的地标节点覆盖面. $\text{Sample}(W, s)$ 表示从集合 W 中选取 s 个节点;当 $s < |W|$ 时,从 W 集合中随机选取 s 个节点,反之,返回整个 W 集合. $\text{RandomOneMax}(k)$ 表示从非地标节点中随机选取一个邻域最大同时大于 k 的节点,若节点邻域都小于 k ,则返回空集.

对于地标节点集合 L ,改进的算法先选取而后再筛选,具体过程如下:首先,选取地标节点集合并计算所有节点的地标节点;然后,统计每个地标节点的覆盖面,并剔除覆盖面小于 k 的地标节点,同时更新节点的地标节点(更新时地标节点的覆盖面都将保持在 k 值以上);最后,获取节点的本地邻域,为保持与 TZ 算法一致,当地标节点的数目小于 s 同时集合 W 为空时,需从非地标节点中随机选取一个邻域规模最大同时大于 k 的节点来更新集合 W ;重复算法直至集合 W 为空或剔除的地标节点集合 D 与添加的地标节点集合 W 一致.

已有的研究表明,在实际的 Internet AS 图上, TZ 算法的迭代次数远小于理论上限 $2\log n$ ^[6].对于改进的算法,其迭代次数与 k 值相关,且当删除的地标节点集合 D 与添加的地标节点集合 W 保持一致时,迭代次数有可能大于原始的上限,但改进算法的地标节点规模却小于原始的规模.

2 实验结果

仿真实验是基于 Internet AS 图来实现的,实验数据来源于 CAIDA^[14],为 2004 年至 2013 年 AS 层的拓扑连接. 对于一个大规模的网络,计算消耗较大,仿真 TZ 算法在计算上是不现实的^[9],因此对于 TZ 算法和 TZ_k 算法,都需要对网络预先处理. 由于 Internet AS 图具有自相似性,用 *k*-shell 法处理网络可以保持网络原始的统计特性^[15]. 为此,本文用 *k*-shell 法提取出核数为 3 的子图用于仿真. 图 2 描述了仿真数据的 Internet AS 图大小,节点数目从 2004 年 11 月的 18 626 个递增到 2013 年 11 月的 45 427 个,其核数为 3 的子图从 4 020 个节点递增到 11 486 个节点.

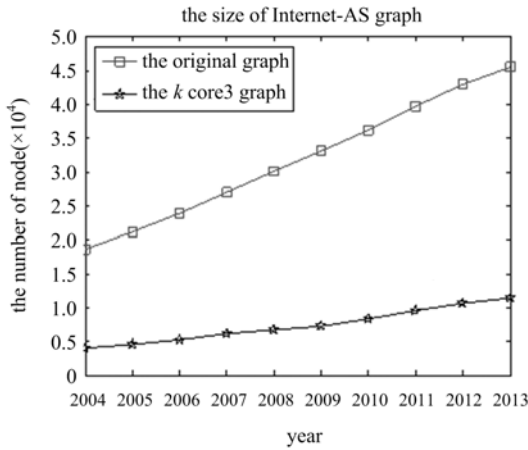


图 2 不同年份的节点数目

Fig. 2 The total nodes with different years

TZ 算法中节点的传输路径是不对称的,故对于每一个节点,仿真其传输包信息到其他所有节点. 初始的地标节点是随机选取的,因此算法存在不确定性,对于不同的地标节点集合,伸长系数、路由表大小都存在差异性. 为此,对于每一个子图,本文仿真 50 次并取平均值作为性能评价.

2.1 地标节点覆盖范围的影响

k 值限定了地标节点的覆盖面,根据 *k* 值的变化,本文研究覆盖面对平均伸长系数、平均路由表及地标节点分布的影响. *k*=0 时,算法相当于原始的 TZ 算法,对选取的地标节点不做任何处理;*k* 越大,迭代过程中剔除的地标节点越多,满足的地标节点的集合越小,当 *k* 大到一定程度,地标节点就只有一

个. 对于图例的分析,文中仅显示 4 年的 AS 图例,当然,其余年份也保持相似的特性.

图 3 展示了不同 *k* 值下,平均伸长系数的分布,其中横轴表示 *k* 值,纵轴表示平均伸长系数,水平的横线表示 TZ 算法的平均伸长系数. 从图 3 可知,随着 *k* 值的增加,平均伸长系数先减小,随后逐渐的增大,且在很长的一段区间内,改进算法的平均伸长系数都比 TZ 算法的小,而且该区间随着网络规模的增加而扩大. 从初始时的下降趋势可知,覆盖面较小的地标节点不利于紧凑路由的性能,适当扩大地标节点的覆盖面可以提高紧凑路由的性能. 随着 *k* 值的增大,地标节点数目减少,经过地标节点中转信息距离变大,因此伸长系数成增大趋势. 当 *k* 值较大时,曲线有稍微的波动,这是因为初始地标节点是随机选取的,而 Internet AS 图具有幂率分布、富俱乐部等特性^[13],度小的节点较多,有更大的概率成为地标节点;度大的节点较少,但其有更好的连通性,当其作为地标节点时,覆盖面往往会很大,因此,初始随机选取地标节点传递信息的差异性在图中呈现出波动性.

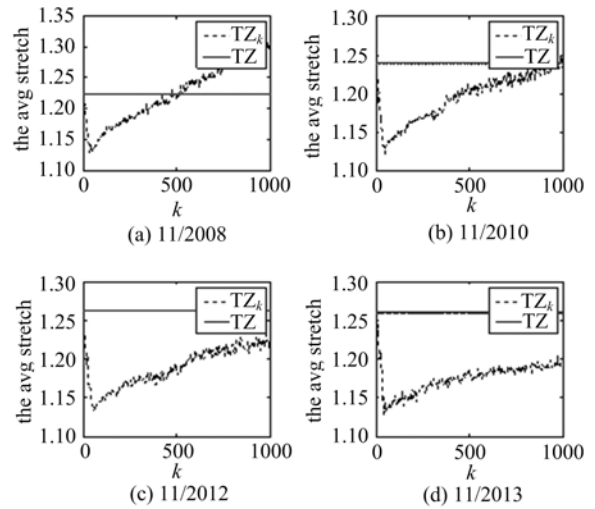
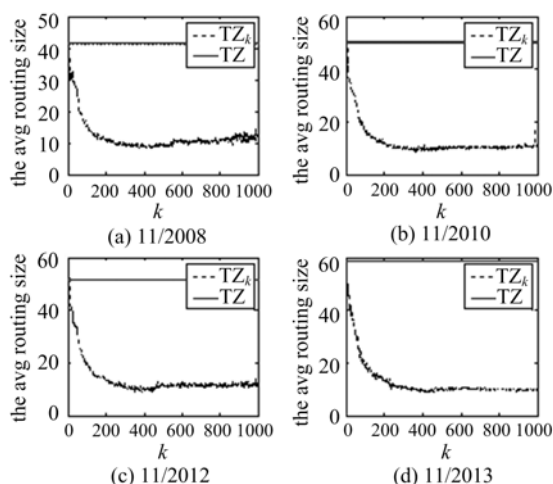


图 3 不同 *k* 值的平均伸长系数分布

Fig. 3 The average stretch distribution with different *k*

图 4 展示了不同 *k* 值下,平均路由表的分布,其中横轴为 *k* 值,纵轴为平均路由表;水平的横线表示 TZ 算法的平均路由表. 从图 4 可知,随着 *k* 值的增加,平均路由表逐渐减小,最后将保持不变.

节点路由表由本地邻域和全局地标节点构成,实际仿真时,本地邻域和全局地标节点规模都远小

图 4 不同 k 值的平均路由表大小分布Fig. 4 The average routing table size distribution with different k

于理论上限^[9]. k 值限制了地标节点的覆盖面, 当 k 值较小时, k 值的变化对地标节点的选取影响较大, 对本地邻域影响不大. 因此, 初始时, 随着 k 的增大, 平均路由表逐步减小. 随着 k 的继续增大, 地标节点数目继续减小, 对本地邻域影响较大, 使其逐步变大, 致使路由表保持不变.

图 5 展示了不同 k 值时, 地标节点的分布特性, 其中横轴为地标节点的覆盖面, 纵轴为节点数目. TZ 算法中地标节点的数目至少为 s 个^[6], 为保持与 TZ 算法一致, 本文仅研究 $|L| \geq s$ 时的分布情况, 由图 3、图 4 可知, 此时平均伸长系数与平均路由表均成递减趋势. 由图 5 可知, $[0; 100]$ 区间的地标节点数目显著降低, $[100; 1000]$ 区间的地标节点数目变化不大, 但总的地标节点数目减少, 可知 k 值很好地约束了地标节点的分布.

k 值限定了地标节点的覆盖面. 针对不同 k 值, 上文分析了平均伸长系数、平均路由表大小的变化趋势, 并发现覆盖面过小的地标节点不利于紧凑路由的效率. 对于紧凑路由, 降低路由表, 就会增大伸长系数, 而降低伸长系数, 也会增大路由表, 因此对于参数 k 的选择, 可根据实际模型来调整. 为尽可能降低传输路径距离, 本文取 $k=s$ 这一伸长系数近似最优的值来限制地标节点的覆盖面并改进 TZ 算法, 同时系统地分析平均伸长系数、平均路由表大小.

2.2 伸长系数

为了更好的分析伸长系数, 本章节详细分析了 2004 年至 2013 年的传输路径距离分布情况.

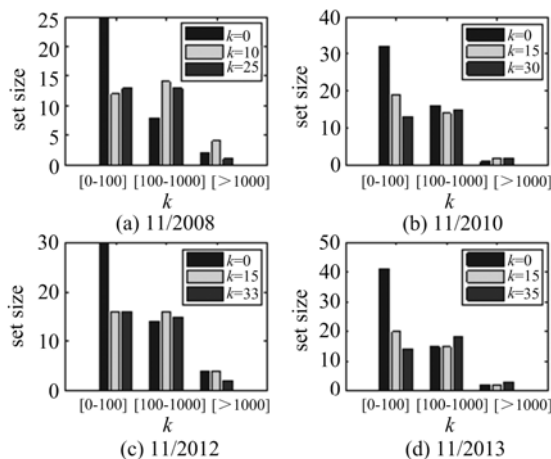
图 5 不同 k 值的地标节点覆盖面的分布Fig. 5 The coverage size distribution of landmark with different k

图 6(a) 展示了 TZ 和 TZ_k 算法的倍数型伸长系数分布. 很明显, TZ_k 算法有效地提高了最短路径所占的比例, 降低了 $[1.5; 2.0]$ 区间的比例. 对于 TZ 算法, 以 2013 年的子图为例, 有 62.62% 的路径其伸长系数为 1; 有 94.41% 的路径其伸长系数小于 1.5; 有 5.35% 的路径伸长系数大于 1.5 小于 2.0. 对于伸长系数处于 $[2.0; 3.0]$ 之间的路径所占的比例仅仅是 0.24%.

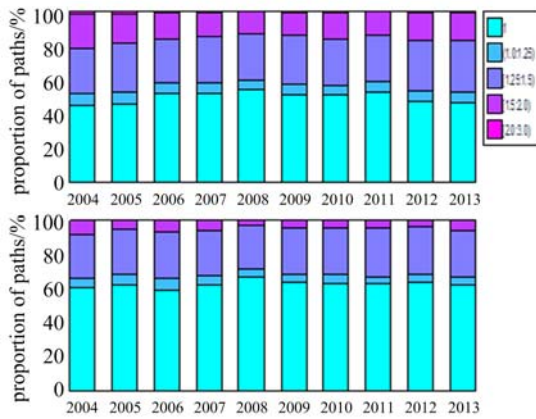
图 6(b) 展示了 TZ 和 TZ_k 的增量型伸长系数. TZ_k 算法有效地提高了最短路径的比例, 降低了附加跳数为 3 的路径比例. 对于 TZ_k 算法, 同样以 2013 年的子图为例, 其中, 31.64% 的路径需要附加 1 跳; 5.8% 的路径需要附加 2 跳数; 仅仅有 0.28% 的路径需要附加多于 2 跳数, 而多于 5 个跳数的路径非常少.

对于伸长系数, 理论上限为 3, 但实际的仿真结果远小于理论上限^[9], 实验结果也证明了这一点. 随着网络规模的递增, 伸长系数依然很小, 而且几乎不存在非常长的路径, 传输路径都近似于最短路径. 已有的研究表明, 对于无标度网络, 随着网络规模的递增, 平均伸长系数有略微的递减^[10], 本文的实验结果也很符合这一特性.

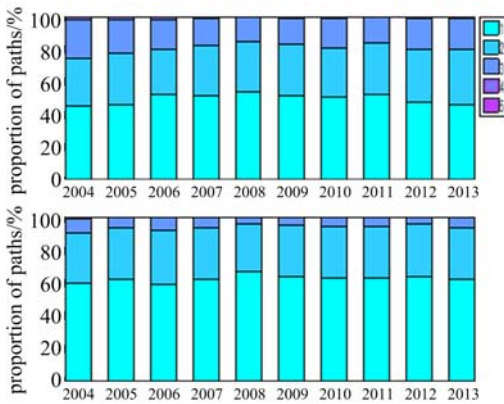
对于 BGP 的 AS 图数据, 多数情况下都是不完整的, 主要是缺少点对点之间的连边^[16], 但增加缺失的连边只会有更好的性能指标, 因此, 相信本文所做出评价的都是最坏情况下的性能.

2.3 路由表大小

图 7 展示了 TZ 和 TZ_k 的路由表大小分布, 很明显 TZ_k 算法有更小的路由表, 而且随着网络规模



(a) Multiplicative stretches for TZ and TZ_k (bottom)



(b) Additive stretches for TZ (top) and TZ_k (bottom)

图 6 TZ 和 TZ_k 的伸长系数

Fig. 6 Path stretch for TZ and TZ_k algorithms

的递增,两者的差距越来越明显.对于 2004 年,地标节点数目从原始的 29 减少到 22,其平均路由表降低了 26.6%;对于 2013 年,地标节点数目从原始的 54 减少到 36,其平均路由表降低了 35.3%.同样,其他年份也有类似的性能,而且路由表依然保持很小.

3 结论

紧凑路由能够解决现在的 BGP 路由扩展性问题.针对 TZ 紧凑路由算法中地标节点覆盖范围的分布特性,本文发现覆盖面过小的地标节点不利于紧凑路由的效率,为此,设定一个阈值 k 来限定地标节点的覆盖面,并在迭代过程中剔除覆盖面小于 k 的地标节点,改进了原始的 TZ 算法,同时系统地分析了地标节点的覆盖面对紧凑路由性能的影响.通过仿真发现,随着地标节点最小覆盖面 k 的增大,平均伸长系数先减小而后逐渐增大,平均路由表先逐

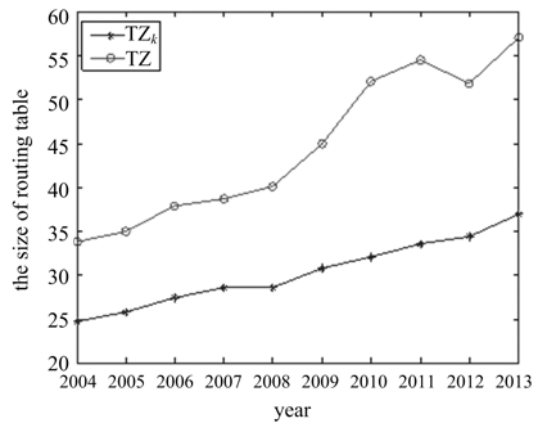


图 7 TZ 和 TZ_k 的路由表大小

Fig. 7 Routing table size for TZ and TZ_k algorithms

步减小而后保持不变;当选取合适的阈值时,改进的算法有更小的平均伸长系数和平均路由表大小,可有效提升 TZ 算法的路由性能.

目前的紧凑路由算法只考虑静态网络,并不适合于动态网络,而且算法与网络的全局拓扑密切相关.所以,针对动态网络,并与网络拓扑无关的紧凑路由算法将是未来一个重要的研究方向.

参考文献 (References)

[1]Edwards B, Hofmeyr S, Stelle G, et al. Internet topology over time [EB/OL]. <http://arxiv.org/pdf/1202.3993v1.pdf>.

[2]Meyer D, Zhang L, Fall K. Report from the IAB works-hop on routing and addressing[EB/OL]. <http://tools.ietf.org/pdf/rfc4984.pdf>.

[3]Li T. Recommendation for a routing architecture[EB/OL]. <http://ftp.fi.netbsd.org/rfc/rfc6115.txt.pdf>.

[4]Krioukov, Fall K, Brady A. On compact routing for Internet [J]. ACM SIGCOMM Computer Communication Review, 2007, 37(3): 41 -52.

[5]Cowen L J. Compact routing with minimum stretch[J]. Journal of Algorithms, 2001, 38(1): 170 -183.

[6]Thorup M, Zwick U. Compact routing schemes[C]// Proceedings of the 13th Annual ACM Symposium on Parallel Algorithms and Architectures. Crete, Greece: ACM Press, 2001: 1 -10.

[7]Faloutsos M, Faloutsos P, Faloutsos C. On power-law relationships of the internet topology [J]. ACM SIGCOMM Computer Communication Review, 1999, 29 (4): 251 -262

- [8] Brady A, Cowen L J. Compact Routing on power law graphs with additive stretch[C]// Proceedings of the 8th Workshop on Algorithm Engineering and Experiments. 2006; 119-128.
- [9] Strowes S D, Mooney G, Perkins C S. Compact routing on the Internet AS-graph[C]// IEEE Conference on Computer Communications Workshops. Shanghai, China; IEEE Press, 2011; 852-857.
- [10] Krioukov D, Fall K, Yang X W. Compact routing on Internet-like graphs [C]// 23rd Annual Joint Conference of the IEEE Computer and Communications Societies. Hongkong, China: IEEE Press, 2004, 1: 209-219.
- [11] 唐明董, 张国清, 杨景, 等. 针对无标度网络的紧凑路由方法[J]. 软件学报, 2010, 21(7): 1732-1743.
TANG Mingdong, HANG Guoqing, YANG Jing, et al. Compact routing scheme for scale-free networks [J]. Journal of Software, 2010, 21(7): 1732-1743.
- [12] Strowes S D, Perkins C S. Harnessing Internet topological stability in Thorup-Zwick compact routing [C]// Proceedings IEEE INFOCOM. Orlando, USA: IEEE Press, 2012; 2551-2555.
- [13] 范涛, 刘勇, 蔡世民, 等. 基于 Internet AS 图的紧凑路由算法研究[J]. 中国科学技术大学学报, 2013, 43(1): 73-78.
FAN Tao, LIU Yong, CAI Shimin, et al. Study on compact routing schemes based on Internet AS-graphs [J]. Journal of University of Science and Technology of China, 2013, 43(1): 73-78.
- [14] Caida. AS Relationships[EB/OL]. <http://www.caida.org/data/active/as-relationships/>.
- [15] Gaertler M, Patrignani M. Dynamic analysis of the autonomous system graph[C]// International Workshop on Inter-domain Performance and Simulation. Budapest, Hungary: IEEE Press, 2004: 13-24.
- [16] Oliveira R, Pei D, Willinger W, et al. The (in) completeness of the observed internet AS-level structure [J]. IEEE/ACM Transactions on Networking, 2010, 18(1): 109-122.