

R-Vivaldi: 距离范围感知的IP网络坐标系统

王大彬^{1,2}, 黄琼¹, 阳小龙^{2,3}, 隆克平³

(1. 重庆邮电大学 移动通信技术重点实验室, 重庆 400065;

2. 电子科技大学 光互联网及移动信息网络研究中心, 四川 成都 611731;

3. 北京科技大学 计算机与通信工程学院, 北京 100083)

摘要: 分析发现, 对需预测的节点间时延, 若选择在该时延范围附近的节点作为锚节点, 则能提高该时延预测准确度。基于此, 提出了一种距离范围感知的IP网络坐标系统, 即R-Vivaldi。其主要思路为: 根据需预测的节点间时延的大致取值范围, 在与该取值范围相近的一个距离半径的空间内, 重新选择锚节点而得到它的一个新取值范围。依照该过程, 被预测时延的取值范围更加明晰, 并不断地动态调整锚节点的选择, 直至节点间时延的预测误差满足一定需要。仿真结果表明, 此方案能有效地提高距离预测准确性, 而且在一定程度上克服了预测准确度不一致性问题。

关键词: IP坐标系统; 网络距离; 距离感知; 坐标计算

中图分类号: TP393

文献标识码: A

文章编号: 1000-436X(2012)02-0087-06

R-Vivaldi: distance range-aware IP network coordinate system

WANG Da-bin^{1, 2}, HUANG Qiong¹, YANG Xiao-long^{2, 3}, LONG Ke-ping³

(1. Key Lab of Mobile Comm. Tech, Chongqing University of Posts and Telecommunications, Chongqing 400065, China;

2. Research Center for Optical Internet and Mobile Info. Netw., Univ. of Electronic Sci. and Tech. of China, Chengdu 611731, China;

3. School of Computer and Communications Engineering, University of Science & Technology Beijing, Beijing 100083, China)

Abstract: The experience showed that the predicted delay's accuracy was improved when the nodes were chosen around the given delay as anchors. Thus a distance range-aware IP network coordinate system(R-Vivaldi) was proposed. The main idea was: according to the approximate range of the predicted delay, the anchors were re-selected in the radius area of the approximate range to compute the new range. In accordance with the above process, the coordinate system constantly iterated the network distance and dynamically chose anchors until the prediction accuracy meet the requirement. Experiments show that the approach can not only effectively improve the prediction accuracy of delay but also overcome inconsistencies of the prediction accuracy to a certain extent.

Key words: IP coordinate system; Internet distance; distance range-aware; coordinate computing

收稿日期: 2010-04-15; 修回日期: 2010-12-13

基金项目: 国家重点基础研究发展计划(“973”计划)基金资助项目(2012CB315905); 国家自然科学基金资助项目(60873263, 60932005, 61172048, 61100184); 四川省青年基金资助项目(09ZQ026-032); 教育部新世纪优秀人才计划基金资助项目(NCET-09-0268)

Foundation Items: The National Basic Research Program of China (973 Program)(2012CB315905); The National Natural Science Foundation of China (60873263, 60932005, 61172048, 61100184); Youth Foundation Program of Sichuan Province(09ZQ026-032); Program for New Century Excellent Talents in University (NCET-09-0268)

1 引言

在实际的网络中, 时延(即网络距离)是一个非常重要的参数, 已把它看作为网络路径的一个基本属性, 与网络拓扑和路由密切相关。如果获得了节点之间的时延信息, 则对提高网络应用(如媒体文件共享、内容访问网络等)的性能有很大的帮助。Ping 方式是获取该信息的最直接方法, 它简单直观, 但是效率低、开销大、可扩展性差, 其时间复杂度为 $O(N^2)$ 。为此, 提出了虚拟坐标系统的概念, 它的基本思想是将网络距离空间映射到一个几何空间中, 每个网络节点对应几何空间中一个坐标点, 节点间距离可以根据它们的坐标值通过空间距离公式计算得出。因此, 虚拟坐标系统能大幅度降低测量开销。

目前, 文献提出了很多不同的虚拟坐标算法, 可以大致划分为基于锚节点(landmark-based)^[1~3]和基于物理模拟(simulation-based)^[4,5]2 类。基于锚节点的坐标系统需要事先配置一定数目的锚节点, 其余节点的坐标通过最小化这些节点到锚节点的实际距离与预测距离的相对误差之和计算得出; 基于物理模拟的坐标系统则是将网络模拟为一种物理模型, 例如 Vivaldi 算法, 它将距离预测误差之和最小化问题模拟为弹簧系统势能最小化问题。尽管这些算法的时延预测相对误差都不大, 但是即使很小的预测误差, 也会对网络应用的性能产生非常明显的影响^[6]。因此, 如何提高它们的时延预测准确度, 成为该领域的一个研究热点。

对该问题目前最常见的解决方案是增加坐标空间维度或增加锚节点数。该方案对维度或者锚节点数较少的坐标系统, 即使维数或锚节点数有少许增加, 其预测准确性就会有明显的改善。但是当坐标维数为 7 以后, 其预测准确性改善就不再明显, 而同时测量开销和坐标计算开销反而会增大。例如: 文献[1]对 GNP(global network positioning)方案的预测误差进行了仿真分析, 其结果为: 坐标系统从 2 维增加到 3 维, 其时延预测的相对误差最大改善可达 5%; 而从 7 维到 8 维、10 维等, 则几乎没有任何改善。鉴于此, 一些研究者则从锚节点的选择入手, 根据被选择的锚节点构造出相应的网络坐标系统。例如文献[7~9]提出的分层选择锚节点和文献[10]提出的混合选择锚节点的方案。其中混合选择锚节点方案是选择一部分邻近节点并在剩余的

节点中, 随机选择一部分节点共同作为锚节点来计算节点的坐标。此方案能提高短距离的预测准确性, 但是同时会降低长距离的预测准确性, 并且对整个坐标系统准确性的改善也不明显; 分层选择锚节点方案如下。1) 按照一定的规则将节点划分到不同的层中, 然后节点选择自己所在层内的节点作为锚节点计算自己在该层的坐标。2) 在每次预测节点间的距离时, 选择相对“正确”的层坐标来预测节点间的距离。此分层选取锚节点方案的优势在于: 不仅能提高短距离的预测准确度, 而且对长距离的预测准确度也不会产生影响。同时还可以从文献[7]的仿真结果中得出。分层选择锚节点的方案比混合选择锚节点的方案具有更高的预测准确性。但是, 文献[7,8]所提出的方案需要一些固定的锚节点来形成簇结构, 所以坐标系统的准确性受到所选择固定锚节点影响; 而文献[9]提出的分层方案, 在网络拓扑改变时, 健壮性很差。

由上述可知分层选择锚节点机制在提高距离预测的准确度上, 优于其他机制, 而且经过研究发现, 该机制在一定程度上克服了预测准确度不一致性问题, 此问题是指对于长距离和短距离节点坐标预测准确度不能到达一致, 从而不能达到全局最优的问题。但该机制也存在很大的局限性, 为了使距离预测准确性不受固定锚节点和网络拓扑改变的影响, 且进一步提高距离预测的准确性, 本文对分层选择锚节点机制进行了扩展, 提出了一种距离范围感知的 IP 网络坐标系统。它是根据虚拟坐标系统中的一种现象而提出的, 即对被预测的节点间时延, 若选择在该时延范围附近的节点作为锚节点, 则能提高该时延预测准确度。仿真结果表明, 与 Vivaldi 算法相比, 此方法有效地提高了距离预测的准确性, 而且在一定程度上克服预测准确度不一致性问题。

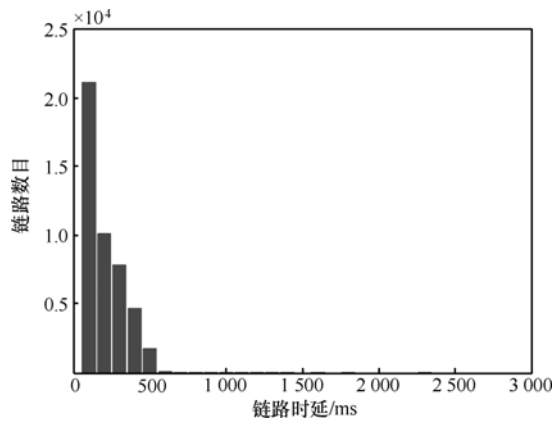
2 R-Vivaldi: 距离范围感知的 IP 网络坐标系统

2.1 问题描述

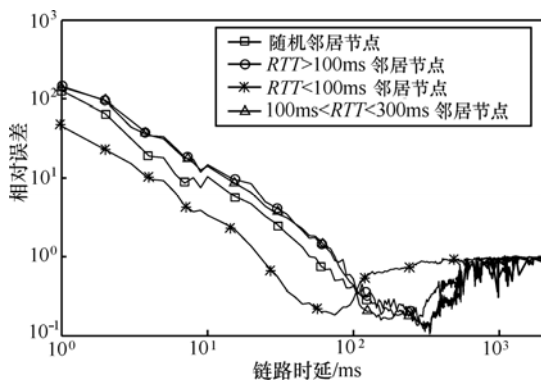
从文献[7~10]可知, 选择不同距离范围内的节点作为锚节点, 对坐标系统的准确性影响很大。本文选取了 Planetlab 的一组真实数据^[11], 以当前最典型的网络坐标 Vivaldi 算法为例, 对该结论进行验证分析。该组数据包括了 226 个主机节点, 将其中在 RTT(round-trip time)小于 100ms 范围内找不到一

定数目邻居节点的节点剔除。在选择邻居节点时, 邻居节点数不宜过多或过少, 前者会导致更多的节点在指定的距离范围内找不到规定的节点数目, 从而使结论失去一般性; 后者会使得计算出的坐标不够准确, 导致结论不够准确。本文仿真中折衷选择 16 个邻居节点。

根据选择不同距离范围内的节点作为邻居节点, 仿真的结果如图 1 所示。



(a) 节点间链路时延分布



(b) Vivaldi 邻居节点的选择范围对时延预测相对误差的影响分布

图 1 不同距离范围内的邻居节点对预测准确性的影响

图 1(a)描述的是节点间链路时延分布情况, 可以看出节点间的时延基本上是小于是 500ms 的。从图 1(b)描述的相对误差分布情况来看, 对于随机选择邻居节点的 Vivaldi 而言, 若选择的邻居节点 $RTT < 100ms$ 时, 提高了时延小于 100ms 的链路预测的准确性, 但是在此值之后随着链路时延的增大, 相对误差急剧上升, 即劣化了对长距离预测的准确性。而当选择的邻居节点 $RTT > 100ms$ 时, 劣化了时延小于 290ms 的链路预测的准确性, 但是在此值之后随着链路时延的增大, 提高了长距离的预测准确度。当选择的邻居节点 $100ms < RTT < 300ms$ 时,

提高了距离在 100ms~270ms 之间的链路预测的准确性。所以得出: 对被预测的节点间时延, 选择在该时延范围附近的节点作为锚节点有利于提高该时延的预测准确性。

但是, 怎么知道被预测时延的大致范围, 以便能取得在该时延范围附近的节点作为锚节点呢? 另外一个问题是如何使锚节点的可选择空间随被预测时延的取值范围动态调整呢? 针对这 2 个问题, 本文提出了 R-Vivaldi 系统。

2.2 算法思想

由上可知, 选择不同距离范围的节点作为锚节点, 对坐标系统的准确性影响很大。若在较小距离范围内选择锚节点, 则可提高相互邻近的节点间时延的预测准确度; 若在较大距离范围内选择锚节点, 则可提高相距较远的节点间时延的预测准确度。总之, 只要能选取在被预测时延范围附近的节点作为锚节点, 就能有效地提高被预测时延的准确性。由此现象可知: 提高节点间时延的预测准确度的关键在于, 如何知道被预测时延的大致范围, 以便能取得在该时延范围附近的节点作为锚节点和如何使锚节点的可选择空间随被预测时延的取值范围动态调整。为此, 本文提出了一种距离范围感知的 IP 网络坐标系。其主要思路为: 根据被预测的节点间时延的大致取值范围, 在与该取值范围相近的一个距离半径的空间内, 重新选择锚节点而得到它的一个新取值范围。依照该过程, 被预测时延的取值范围更加明晰, 并不断地动态调整锚节点的选择, 直至节点间时延的预测误差满足一定要求。

2.3 R-Vivaldi 的实现

该系统的实现过程如图 2 所示, 且以预测节点 A 与 B 的距离为例。首先, 节点被嵌入到某个几何空间中并拥有了一个几何坐标, 记该坐标为全局坐标, 此时的状态如图 2(a)所示; 接下来, 如图 2(b)所示, 根据 2 个节点的全局坐标计算 2 个节点间的距离 D_{AB} , 然后在以节点 A (或 B) 为中心的一个环形区域内选取其中节点作为锚节点 (标记为 1 的节点), 计算 A (或 B) 的第 1 层坐标。该环形区域是以节点 A (或 B) 为中心, 半径为 $(1 + \alpha) \times D_{AB}$ (其中 α 是调节参数) 的圆形区域减去以节点 A (或 B) 为中心, 半径为 $(1 - \alpha) \times D_{AB}$ 的圆形区域所得。依此类推, 如图 2(c)所示, 根据两节点第 $n-1$ 层坐标更新节点间的距离 D_{AB} , 然后再以节点 A (或 B) 为中心的一个环形区域内选取其中节点作为锚节

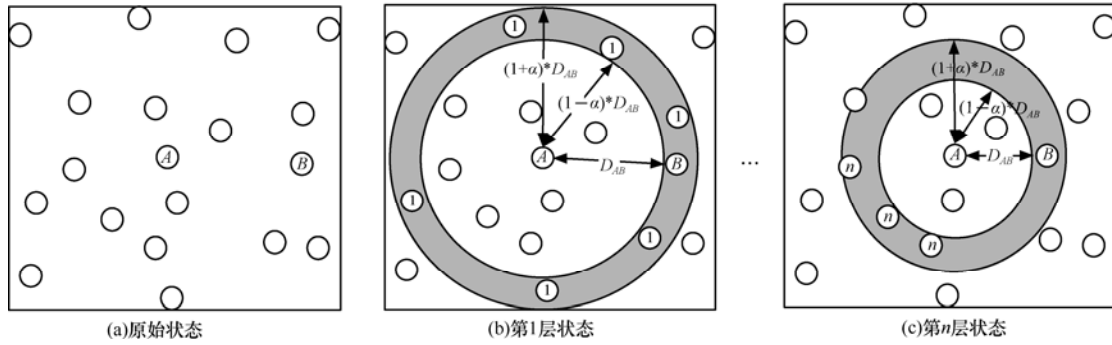


图 2 R-Vivaldi 的实现过程

点 (标记为 n 的节点), 计算 A (或 B) 的第 n 层坐标, 该环形区域定义同上。依照该过程, 使节点 A 与 B 的时延取值范围更加明晰。该过程在预测距离稳定即预测距离的准确度达到一定程度或者出现异常时停止, 并利用最后一层的坐标预测两节点间的距离。

R-Vivaldi 伪代码

/*初始化, 其中 N_C 与 N_A 分别是迭代次数和锚节点数目, α 和 β 是调节参数, C_A 和 C_B 分别是节点 A 与 B 的坐标*/

$T_A=C_A, T_B=C_B, N_C=n, N_A=m, \alpha=\alpha', \beta=\beta'$

//计算两节点 A 与 B 之间的距离

$D_{AB}=\|C_A-C_B\|$

//在 $[(1-\alpha) \cdot D_{AB}, (1+\alpha) \cdot D_{AB}]$ 内, 选择新锚节点

while $N_C > 0$ and $|D_{AB} - D'_{AB}| / D'_{AB} > \beta$
do

for node i in the network do

$L(A-i)=\text{Measure Distance from } A \text{ to } i$

if $L(A-i) \geq (1-\alpha) \cdot D_{AB}$ and $L(A-i)$

$\leq (1+\alpha) \cdot D_{AB}$ then

$\text{Join_Anchor}(A)$

end if

end for

for node j in the network do

$L(B-j)=\text{Measure Distance from } B \text{ to } j$

if $L(B-j) \geq (1-\alpha) \cdot D_{AB}$ and $L(B-j) \leq (1+\alpha) \cdot D_{AB}$ then

$\text{Join_Anchor}(B)$

end if

end for

//如果在 $[(1-\alpha) \cdot D_{AB}, (1+\alpha) \cdot D_{AB}]$ 内, 选择的锚节点的数目达不到 N_A , 算法停止

if $| \text{Join_Anchor}(A) | < N_A$ or $| \text{Join_}$

$\text{Anchor}(B) | < N_A$ then

end while

end if

for $c=1: N_A$ do

$x = \text{random}(x \text{ in } \text{Join_Anchor}(A))$

$C'_A = \text{Update_coordinate_algorithm}(x)$

end for

for $c=1: N_A$ do

$y = \text{random}(y \text{ in } \text{Join_Anchor}(B))$

$C'_B = \text{Update_coordinate_algorithm}(y)$

end for

$D'_{AB} = \|C'_A - C'_B\|$

$\text{wait}(\text{Update_} C'_A, C'_B, D_{AB}, N_C)$

end while

其中, α 和 β 是调节参数且它们的取值都有一定的要求。 α 取值决定了环形区域的宽度, 所以 α 值过大, 环形区域也就越大, 导致算法收敛速度慢; 而过小, 环形区域越窄, 易导致程序因为节点在该环形区域内找不到足够的满足条件的锚节点而停止, 使得结果不够准确。 β 取值决定了坐标准确度应该满足的条件, 它的值太大易使结果不准确, 而太小使得程序收敛速度慢。

3 算法仿真结果及分析

3.1 算法性能评价标准

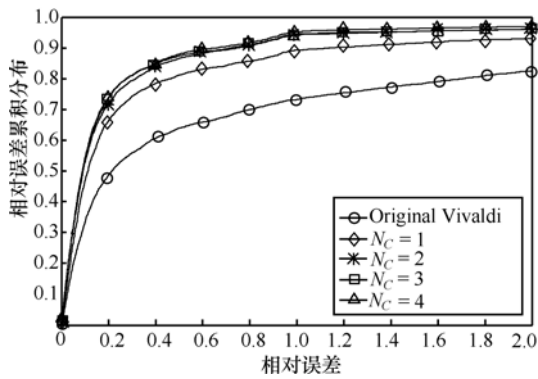
评价一个虚拟坐标系统的准确性总是使用相对误差, 它反映了预测距离与测量时延之间的误差, 定义如下:

$$\text{相对误差} = \left| \frac{\text{预测距离值} - \text{测量距离值}}{\text{测量距离值}} \right|$$

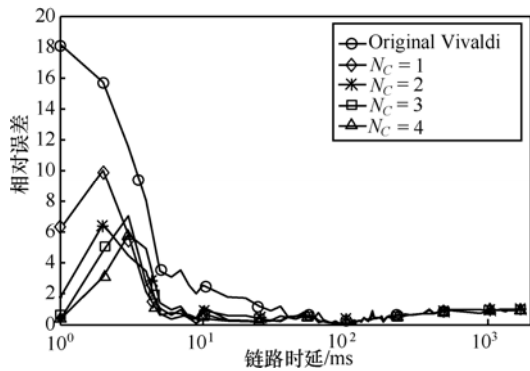
3.2 仿真结果及分析

此处, 仿真数据来源于网络中测量的真实数据

——Planetlab 数据, 它包括了 226 个主机。仿真中, 随机选择了其中 80 个节点, 组成了 80×80 的时延矩阵, 邻居节点的数目设为 16, 坐标维度为 5 维, 调节参数 α 的取值分别为 0.1、0.3、0.5、0.7; 调节参数 β 的取值, 本处为了仿真的方便, 设它的值为 0; 该算法迭代的次数 N_c , 取值分别为 1、2、3、4。仿真结果如图 3 和图 4 所示。



(a) 相对误差累积分布

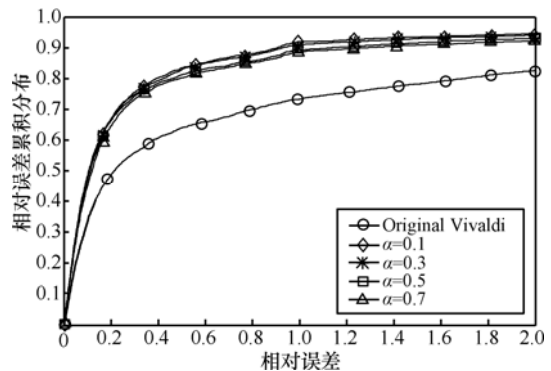


(b) 相对误差分布

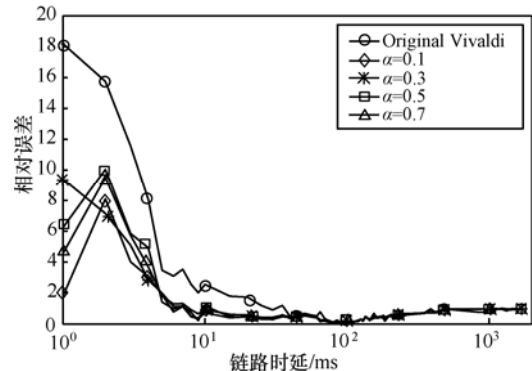
图 3 不同的 N_c 对预测准确性的影响 (α 的值为 0.5)

从图 3(a)和图 4(a)可以看出, 本文提出的 R-Vivaldi 在预测距离准确性上大大优于 Vivaldi 算法, 而且, 在图 3(a)中, 还可以看出, 随着迭代次数的增加, 预测准确度也在提高; 但是当迭代的次数达到一定的数目后, 再增加迭代次数, 预测的准确性也不会再有多大的改善。图 4(a)描述了不同的 α 值对预测准确度的影响: α 值越小, 预测准确度越高。从图 3(a)中, 可以看到在 Vivaldi 算法中, 相对误差为 0.5 和 1.0 时, 相对误差累积分布值分别为 56% 和 74%, 而 R-Vivaldi 相对应的值为 82% 和 96% (在 $N_c=4$ 时)。在图 4(a)中, 在相对误差为 0.5 和 1.0 时, 相对误差累积分布的值为 84% 和 92% (在 $\alpha=0.1$ 时)。

图 3(b)和图 4(b)描述了相对误差的分布情况, 可从图中看出被预测的链路时延小于 100ms 时, R-Vivaldi 大幅度地减少了相对误差值, 而且如图 3(b)所示, 在链路时延大于 5ms 时, 整个坐标系统的相对误差都是小于 1 的。



(a) 相对误差累积分布



(b) 相对误差分布

图 4 不同的 α 值对预测准确性的影响 (N_c 的值为 1)

4 结束语

在本文中, 为了提高网络距离预测的准确性和克服现有的一些分层选择锚节点机制的不足, 提出了一种距离范围感知的 IP 网络坐标系统。对被预测的节点间时延, 选择在该时延范围附近的节点作为锚节点, 能有效地提高该时延预测的准确性。仿真结果显示, 本方案能有效提高距离预测准确性, 而且在一定程度上克服预测准确度不一致性问题。但是, 该系统对长距离的预测相对误差并没有改善, 而且短距离的预测相对误差还是相对较高, 如图 3(b)所示, 在链路时延小于 5ms 时, 相对误差还是比较高。所以, 在以后的工作中, 继续完善本系统, 使它进一步提高距离预测的准确性。

参考文献:

- [1] NG T, ZHANG H. Predicting Internet network distances with coordinate-based approaches[A]. Proc of IEEE INFOCOM'02[C]. New York, 2002.170-179.
- [2] LEE S, SAHU S. A cluster based approach for network distance embedding[A]. Proc of the 9th International Conference on Communications and Information Technologies[C]. Icheon, 2009.1040-1045.
- [3] ZHANG Y, ZHANG H. A steady network coordinate system for network distance estimating[A]. Proc of the 2009 15th International Conference on Parallel and Distributed Systems[C]. Shenzhen, China, 2009. 860-863
- [4] DABEK F, COX R, KAASHOEK F, *et al.* Vivaldi: a decentralized network coordinate system[A]. Proc of ACM SIGCOMM'04[C]. Portland, OR, 2004.15-26.
- [5] SHAVITT Y, TANKEL T. Big-bang simulation for embedding network distances in euclidean space[J]. IEEE/ACM Transactions on Networking, 2004, 12 (6): 993-1006.
- [6] ZHANG R, TANG C, HU Y. Impact of the inaccuracy of distance prediction algorithms on Internet applications—an analytical and comparative study[A]. Proc of the IEEE INFOCOM[C]. Barcelona, 2006. 1-12.
- [7] ZHANG R, HU Y, LIN X, *et al.* A hierarchical approach to Internet distance prediction[A]. Proc of ICDCS[C]. Lisboa, Portugal, 2006. 73.
- [8] CHEN Y, XIONG Y, SHI X. Pharos: a decentralized and hierarchical network coordinate system for Internet distance prediction[A]. Proc IEEE GLOBECOM[C]. Washington, DC, 2007.26-30.
- [9] YE Z, LIU Y, CHEN S. A new hierarchical network coordinate algorithm based on community structure[A]. International Conference on Computational Science and Engineering[C]. Vancouver, 2009. 418-423.
- [10] COSTA M, CASTRO M, ROWSTRON A. PIC: practical Internet coordinates for distance estimation[A]. Proc of IEEE ICDCS[C]. Washington, DC, 2004. 178-187.
- [11] <https://www.planet-lab.org/>, 2006[EB/OL].

作者简介:



王大彬 (1985-), 男, 四川宜宾人, 重庆邮电大学硕士生, 主要研究方向为计算机网络。



黄琼 (1971-), 女, 四川西昌人, 重庆邮电大学副教授, 主要研究方向为个人通信、光互联网及交换技术。



阳小龙 (1970-), 男, 四川邻水人, 博士, 北京科技大学教授、博士生导师, 主要研究方向为互联网理论与技术、网络安全、光交换与光互联。



隆克平 (1968-), 男, 四川通江人, 博士, 北京科技大学教授、博士生导师, 主要研究方向为互联网理论与技术、光互联网体系结构和关键技术、宽带无线接入技术、电信增值业务等。