

# T-Vivaldi: TIV感知的IP网络坐标系统

黄琼<sup>1</sup>, 刘熙<sup>1,3</sup>, 阳小龙<sup>2</sup>, 隆克平<sup>2</sup>

(1. 重庆邮电大学通信与信息工程学院 重庆 沙坪坝区 400065;

2. 北京科技大学计算机与通信工程学院 北京 海淀区 100083;

3. 电子科技大学通信与信息工程学院 成都 611731)

**【摘要】**提出了一种对TIV进行检测和抑制的方法,该方法用三角不等式条件检测违例边,并使用违例系数度量其违例程度。根据该系数的值抑制违例边对坐标的更新,从而达到抑制TIV对坐标系统的影响的目的。基于该方法对Vivaldi算法进行了改进,提出了对TIV感知的IP网络坐标系统TIV-aware Vivaldi(T-Vivaldi)。仿真结果表明,与Vivaldi相比,T-Vivaldi能有效地抑制TIV的影响,从而提高预测的准确度。

**关键词** 检测; 网络距离; 抑制; 三角不等式违例

中图分类号 TP393

文献标识码 A

doi:10.3969/j.issn.1001-0548.2012.01.028

## T-Vivaldi: TIV-Aware Coordinates System for IP Networks

HUANG Qiong<sup>1</sup>, LIU Xi<sup>1,3</sup>, YANG Xiao-long<sup>2</sup>, and LONG Ke-ping<sup>2</sup>

(1.College of Communication and Information Engineering, Chongqing University of Posts and Telecommunications Shapingba Chongqing 400065;

2. School of Computer and Communication Engineering, University of Science and Technology Haidian Beijing 100083;

3. School of Communication and Information Engineering, University of Electronics Science and Technology of China Chengdu 611731)

**Abstract** This paper proposes a new coordinates system (T-Vivaldi) which uses a mechanism to detect and restrict triangle inequality violations. The mechanism can detect TIV with triangle inequality qualification, use a parameter to quantify its degree, and restrict TIV through restricting coordinates update with the parameter. The simulation results show that: compared with Vivaldi, T-Vivaldi can restrict the impact of TIV and increase the predicting accuracy.

**Key words** detecting; network distance; restricting; triangle inequality violation

简化而高效的IP网络时延测量方法,是近年来网络测量领域的研究重点之一。一些研究者提出了基于虚拟坐标的间接测量方法,将IP网络时延嵌入到某一测度空间,建立起网络节点与该空间点的映射关系,为网络节点赋予一组相应的空间坐标,由此网络中节点间的时延与空间内两点之间的测度距离产生一定的对应关系,时延测量任务转为空间内两点之间的测度距离计算任务。

文献[1]中,网络距离常用网络时延表示,使用虚拟坐标预测网络距离。目前的虚拟坐标系统主要分为两大类:1)基于锚节点,如GNP<sup>[1]</sup>、PIC<sup>[2]</sup>;2)基于非锚节点,如Vivaldi<sup>[3]</sup>。基于锚节点的坐标系统需要事先配置一些节点(称为锚节点)作为参考节点,这些参考节点使用它们之间的网络时延,再

结合一定的优化算法(如单纯形下降法),以最小化相互间的实际距离与预测距离的误差之和为目标计算自己的坐标。新节点加入系统后,首先测量到参考节点的时延并获取其坐标,然后采用优化算法,以最小化自己到参考节点的实际距离与预测距离的误差之和为目标,获取自己在系统中的坐标。基于非锚节点的坐标系统则是将网络模拟为一种物理模型,例如“弹簧系统”和“力场作用下的粒子运动”。

如何提高IP网络坐标系统的时延预测准确度一直是网络空间嵌入研究领域的热点之一。根据文献[4]可知,由于实际网络中路由策略的影响,节点间时延经常出现三角不等式违例(triangle inequality violations, TIV)现象,而违例会使得节点坐标的计算产生误差而影响系统的时延预测准确性。目前的IP

收稿日期: 2010-03-16; 修回日期: 2010-09-16

基金项目: 国家973计划(2007CB310706, 2012CB315905); 国家自然科学基金(60932005, 61172048, 61100184); 国家科技重大专项(2011ZX03005-004-02); 教育部新世纪优秀人才计划(NCET-09-0268); 四川省青年基金(09ZQ026-032); 重庆市教委科研项目(KJ100514)。

作者简介: 黄琼(1971-),女,副教授,主要从事计算机网络方面的研究。

网络坐标系统很少将TIV考虑在内,而TIV是IP网络中存在的不可避免的固有现象,其对坐标系统的影响不能根除,只能采取相应措施抑制其影响。

本文对TIV带来的不利影响进行分析,提出了TIV检测和抑制的方法。基于此,对Vivaldi算法进行改进,提出了一种TIV感知的IP网络坐标系统T-Vivaldi。T-Vivaldi在对违例边进行检测以后,以违例系数度量其违例程度,并根据该系数的值抑制坐标的更新,从而达到抑制违例所带来的不利影响的目的。仿真结果表明,在不同的违例环境下,T-Vivaldi的预测准确度比没有考虑TIV的Vivaldi有一定程度的提高。

## 1 问题描述

### 1.1 违例边的定义和三角不等式违例的影响

根据文献[2]可知,由于路由策略的影响,实际网络节点之间的时延并不严格满足三角不等式,如图1所示。实际网络中有3个节点A、B、C,用 $L(A,B)$ 、 $L(A,C)$ 、 $L(B,C)$ 表示它们之间的时延,如果 $L(A,B)+L(A,C)<L(B,C)$ ,将边BC称为造成TIV的违例边, $L(B,C)$ 是违例边的时延。

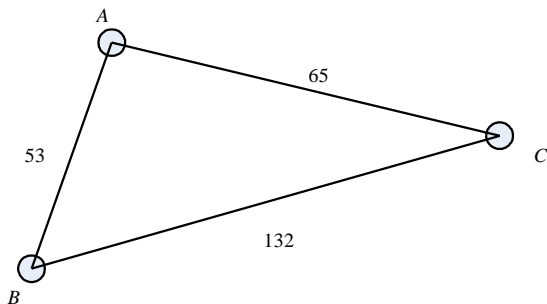


图1 三角不等式违例

TIV是实际网络空间中存在的不可避免的固有现象。TIV的存在表明实际网络空间节点之间的时延不完全满足测度空间特性,因此在进行时延嵌入时会造成误差,该误差是影响预测准确度的主要因素。根据文献[5]可知,在将时延嵌入到测度空间后,产生违例的3个节点之间的违例边会被缩短,而其他两条边会被拉长。为了说明违例对网络坐标系统的预测准确度的影响,本文使用TIV占不同比例的3组时延数据在Vivaldi系统中进行仿真。每组时延数据包含60个节点之间的RTT值,TIV的比例分别是22.47%、31.83%、44.58%。3组比例用违例三角形的数量与所有节点构成的三角形的总数量的比值表示。统计基于3组数据获得的预测距离与实际距离的相对误差的累积分布,如图2所示。

由图2可以看出,随着TIV比例的增加,坐标系

统的预测准确度下降。为了提高预测准确度,必须要抑制违例对坐标系统的影响。

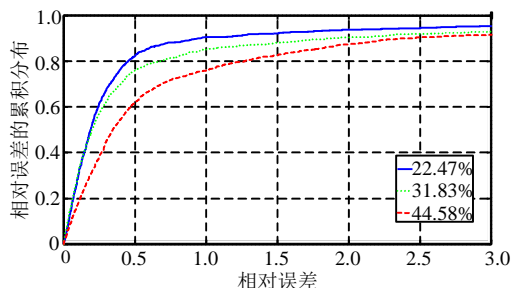


图2 TIV占不同比例下相对误差的累积分布

### 1.2 三角不等式违例的研究现状

针对网络时延中的三角不等式违例现象,不同文献提出了各自的解决方法。文献[6-8]分析了坐标维数与系统准确度的关系,用实际数据指出当坐标维数达到7维以后,继续增加维数已经不能更进一步提升预测准确度。文献[9]分析节点间时延的大小与三角不等式违例的关系,提出了使用分层次坐标预测不同范围内的时延的方法。首先对节点按它们之间的相互距离进行聚类,分为不同的簇,每个簇内节点间的时延不能超过一个预先规定的最大值。簇内的节点使用局部坐标体系,簇间的节点使用全局坐标体系。文献[10]采用矩阵的正交分解方法预测节点时延。将网络节点间的时延矩阵正交分解为两个矩阵的乘积形式,每个节点被赋予了1个出向量和1个入向量、2个向量。被乘数矩阵的行向量表示节点的出向量,乘数矩阵的列向量表示节点的入向量。节点 $i$ 到节点 $j$ 的距离用 $i$ 的出向量和 $j$ 的入向量的点积计算。文献[5]提出了一种对违例边进行预测的机制,使用的基本原理是,因违例边嵌入测度空间后会产生收缩,可根据收缩的程度判断边违例的可能性。收缩越严重,违例产生的可能性就越大。

与没有考虑TIV的坐标系统相比,上述方法都有更高的预测准确度。TIV对坐标系统的影响主要是违例边产生的,而解决的办法也相应针对对违例边的处理。本文针对违例边提出了对TIV进行检测和抑制的方法,并基于该方法对Vivaldi进行改进,提出了T-Vivaldi坐标系统。

## 2 对三角不等式违例感知的网络坐标系统——T-Vivaldi

三角不等式违例在网络中并不遵从特有的分布规律,违例可能出现在网络中任意3个节点之间,因此对违例的检测和抑制要以分布式的方式实现。而Vivaldi是一种完全分布式的坐标系统,本文的方法可以与其很好地结合。

## 2.1 Vivaldi坐标系

Vivaldi系统<sup>[3]</sup>中每个节点要测量自己到一定数量的邻居节点的时延, 并通过最小化网络节点间的实际距离与预测距离的误差逐步更新自己的坐标, 最后达到系统的稳定。

下面是Vivaldi算法中单个节点的坐标计算的具体实现过程。

假设点*i*的坐标是 $x_i$ , 用参数 $e_i$ 表示该点对自己的误差的估计值。它的一个邻居节点为*j*, 要更新*i*的坐标, *i*测量自己到*j*的距离(RTT), 并引入调节参数 $c_e$ 、 $c_c$ 和单位向量 $u(\cdot)$ 。

计算本地坐标与其他坐标的误差权重:

$$\omega = e_i / (e_i + e_j) \quad (1)$$

计算两点之间距离的相对误差(计算值与实测值之间的相对误差):

$$e_s = \|x_i - x_j\| - \text{RTT} / \text{RTT} \quad (2)$$

计算本地误差的变化:

$$e_i = e_s c_e \omega + e_i (1 - c_e \omega) \quad (3)$$

更新本地坐标为:

$$\delta = c_c \omega \quad (4)$$

$$x_i = x_i + \delta (\text{RTT} - \|x_i - x_j\|) u(x_i - x_j) \quad (5)$$

实际上, 在刚开始时, Vivaldi系统中节点坐标的精确度较低; 随着更新的进行, 坐标的精确度逐步提高, 整个系统最后达到一个稳定状态。

## 2.2 T-Vivaldi系统详述

从Vivaldi算法可以看出, 系统本身并没有考虑TIV的情况, 其默认任意节点之间的时延是满足三角不等式关系的, 因而违例会影响坐标的计算, 需要将违例的抑制方法加入到系统中。

文献[5]指出三角不等式违例对坐标系统的不利影响主要通过违例边产生。在Vivaldi系统中, 违例边的影响表现在坐标更新的过程中。违例将导致网络坐标系统在一些区域出现振荡, 会降低系统的稳定性, 而且出现振荡区域的大小与违例程度密切相关。如果节点使用违例边的时延更新自己的坐标, 相比于不违例的边, 坐标的更新幅度会较大, 违例越严重, 更新的幅度越大, 偏离准确值的程度也越大。因此, 可以用抑制违例边的时延对节点坐标的更新抑制TIV的不利影响。

下面是T-Vivaldi对TIV的处理方法。

1) 对TIV进行感知。T-Vivaldi中的节点在更新自己的坐标时, 仍然要测量其与邻居节点的时延, 并使用该时延更新自己的坐标。而邻居节点的选取是随机的, 该点到邻居节点的时延是否造成TIV并不

清楚, 所以要先对该时延进行TIV的判断, 即对其是否违例进行感知。T-Vivaldi使用邻居节点间的时延并结合三角不等式关系进行上述过程。

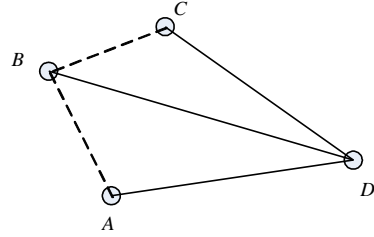


图3 违例边的检测

如图3所示, *D*是需要更新坐标的节点。当*D*开始坐标更新时, 先要测量其邻居节点的时延RTT, 假设其中的3个邻居节点*A*、*B*、*C*, 用 $L(D, C)$ 、 $L(B, C)$ 、 $L(B, D)$ 表示*B*、*C*、*D*相互之间的时延。如果以 $L(D, C)$ 更新*D*的坐标, 首先要对边*DC*是否违例进行判断, 并使用一个违例系数 $\lambda$ 定量边*DC*的违例程度。使用节点*B*作为参考节点(选取方案见2.3节), 在三角形*BCD*中进行违例的检测。定义边*DC*的违例系数为:

$$\lambda = \frac{L(D, C)}{L(B, C) + L(B, D)} \quad (6)$$

如果 $\lambda \leq 1$ , 边*DC*不是违例边; 如果 $\lambda > 1$ , 边*DC*是违例边。 $\lambda$ 表示*DC*的违例严重度,  $\lambda$ 越大, 说明违例越严重。 $\lambda$ 的值与参考节点的选取密切相关, 其值随着参考节点的变化而变化, 如何选取更有价值的参考节点对 $\lambda$ 的计算有重要的影响, 因此在2.3节中介绍了参考节点的选取方案。

2) 对TIV的抑制。在加入违例系数 $\lambda$ 对违例严重度进行量化后, 根据量化值决定违例边对坐标更新的程度。将式(4)中 $\delta$ 的计算变为:

$$\delta = \varepsilon c_c \omega \quad (7)$$

其中:

$$\varepsilon = \begin{cases} 1 & \lambda \leq 1 \\ \frac{1}{\lambda} & \lambda > 1 \end{cases} \quad (8)$$

式(7)和式(8)的意义在于, 使用 $L(D, C)$ 更新*D*的坐标时, 如果边*DC*没有违例, *D*坐标的更新按照正常的算法进行; 如果边*DC*是违例边, 就要抑制其对坐标的更新, 即减少违例边对坐标更新的贡献度。由此可看出, 违例越严重(即 $\lambda$ 越大),  $\delta$ 越小, 坐标更新的程度也越小, 从而违例边对坐标更新的影响也越小, 就可以抑制TIV对坐标系统的影响。

## 2.3 T-Vivaldi中参考节点的选取方案

本文详细描述参考节点的选取方案。首先给出节点坐标更新度的定义。

**定义 1** 节点坐标更新度指节点坐标在更新前后的变化程度。如在更新前节点的坐标为 $C$ ,更新后节点的坐标为 $C'$ ,那么该节点的坐标更新度可以表示为:

$$C_d = |C' - C| \quad (9)$$

TIV对坐标系统的影响主要是造成节点坐标的大幅度振荡,导致其较大幅度的变动,即该节点的 $C_d$ 值较大。而受TIV影响较小或不受其影响的节点的 $C_d$ 值较小。基于上述现象,本文选择坐标更新度较小的节点作为参考节点,可有效地检测出违例边。

参考节点的选取是动态变化的,当满足以下两个条件时,T-Vivaldi系统会更改参考节点:1)新参考节点所在的边被证明不是违例边;2)新参考节点的坐标更新度小于当前参考节点的坐标更新度。下面举例说明参考节点的更改过程。如图3所示,如果当前的参考节点为 $B$ ,在使用 $DC$ 边更新 $D$ 的坐标时,节点 $C$ 和节点 $B$ 要同时提供它们的坐标更新度,然后根据式(6)计算 $DC$ 的违例程度 $\lambda$ 。如果 $\lambda \leq 1$ ,表明 $DC$ 不是违例边。在该情况下,如果 $C$ 的坐标更新度同时又小于 $B$ 的坐标更新度,那么T-Vivaldi系统就会将 $C$ 提升为新的参考节点。如果不同时满足上述两个条件,参考节点不发生变化。

同时,T-Vivaldi给每个节点都分配了一个参考节点,每个节点独立选取和更改自己的参考节点,这样可以避免因使用同一参考点造成单个节点的负担过大。

### 3 T-Vivaldi系统性能仿真

为了验证T-Vivaldi系统的性能,本文在linux系统下进行仿真。仿真采用的时延数据来自Planetlab,下载自哈佛的网络坐标研究小组<sup>[11]</sup>。Planetlab是全球规模最大的verlay网络应用试验床,能够测量并收集真实网络中主机间的多种性能参数值。选取169个节点之间的RTT,为了说明T-Vivaldi处理违例时的有效性,从169个节点中选取TIV占不同比例的3组时延数据进行仿真。该比例用违例三角形的总数量与所有节点能构成的总的三角形数量的比值表示。每组数据包含60个节点间的时延,TIV的比例分别为22.47%、31.83%、44.58%。

在网络距离预测系统中,准确度是一个重要的参数指标。系统距离预测的准确度利用节点间网络距离的预测值和实际值之间的相对误差评价。相对误差定义为:

$$\text{相对误差} = \frac{|\text{预测距离值} - \text{测量距离值}|}{\min(\text{预测距离值}, \text{测量距离值})} \quad (10)$$

本文分别用Vivaldi和T-Vivaldi对每组时延数据进行仿真。参数设置如下:节点个数为60,邻居个数为16,坐标维数为7维,迭代次数为6 000。

输入数据是节点间的时延,输出结果是每个节点的坐标。统计每组数据下两者的相对误差的累积分布,分别如图4~图6所示。

图中虚线是Vivaldi的相对误差累积分布,实线是T-Vivaldi的相对误差累积分布。在同一图中,实线在上说明距离预测的准确度更高。由图可以看出,在三角不等式违例占22.47%时,T-Vivaldi比Vivaldi有一定的提升,但较为轻微。违例占31.83%时,提升效果已经较为明显;误差为0.5时,T-Vivaldi达到了80%,比Vivaldi的76%提高了4%。在违例为44.58%的时延环境下,T-Vivaldi的提升更大,在误差为0.5时,T-Vivaldi比Vivaldi提高了13%。

从3组实验数据的结果可以看出,T-Vivaldi的预测准确度相比于Vivaldi有一定程度的提高。特别是随着违例比例的增加,提升的效果更明显,说明T-Vivaldi在抑制三角不等式违例的影响时,起到了一定的效果。

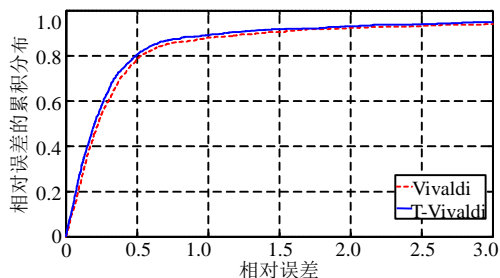


图4 违例占22.47%时的误差累积分布比较

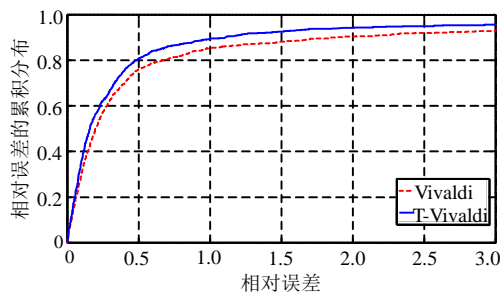


图5 违例占31.83%时的误差累积分布比较

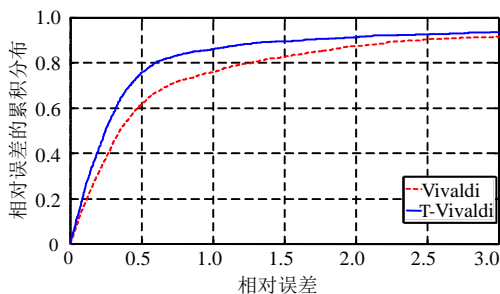


图6 违例占44.58%时的误差累积分布比较

本文使用从planetlab上取得的全部169个节点之间的RTT数据进行仿真, 统计了相对误差的累积分布, 结果如图7所示。仿真参数同上。

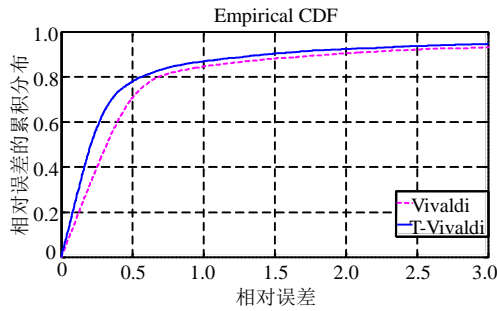


图7 使用planetlab数据时的误差比较

从图7中可以看出, 相对误差在0~0.5之间时, T-Vivaldi比Vivaldi的预测准确度有较明显的提升; 特别是在误差为0.4时, T-Vivaldi的72%比Vivaldi的60%提升了12%, 进一步证明了T-Vivaldi能够起到抑制三角不等式违例的不利影响的作用。

## 4 结论

本文针对网络虚拟坐标系统中的三角不等式违例问题提出一种检测与抑制的机制, 并基于该机制对Vivaldi进行改进, 建立了对三角不等式感知的T-Vivaldi系统。通过不同的数据仿真结果表明, T-Vivaldi能有效地抑制TIV的影响, 从而提高预测的准确度。

### 参考文献

[1] NG E, ZHANG H. Predicting internet network distances with coordinate-based approaches[C]//Proc of IEEE INFOCOM'02. New York, NY: IEEE, 2002.

[2] ZHENG H, LUA E K, PIAS M, et al. Internet routing policies and round-trip-times[C]//Proc of PAM'2005. Boston, MA, USA: ACM, 2005.

[3] DABEK F, COX R, KAASHOE F, et al. Vivaldi: a decentralized network coordinate system[C]//Proc of ACM SIGCOMM'04. Portland, OR, USA: ACM, 2004.

[4] COSTA M, CASTRO M, ROWSTRON A, et al: PIC: Practical internet coordinates for distance estimation[C]//International Conference on Distributed Systems. Tokyo, Japan: [s.n.], 2004.

[5] WANG G, ZHANG B, NG T S E. Towards network triangle inequality violation aware distributed systems[C]//Proc of the SIGCOMM. New York, NY, USA: [s.n.], 2007.

[6] WONG B, SLIVKINS A, SIRER E. Meridian: a lightweight network location service without virtual coordinates[C]//Proc of the ACM SIGCOMM. New York, NY, USA: ACM, 2005.

[7] LEE Sanghwan, ZHANG Zhi-li, SAHU S, et al. On suitability of euclidean embedding of Internet hosts[C]//Proc of SIGMETRICS 2006. Saint-Malo, France: ACM, 2006.

[8] LUA E K, GRIFFIN T, PIAS M, et al. On the accuracy of embeddings for internet coordinate systems[J]. IEEE Transactions on Networking, 2010, 18(1): 27-40.

[9] KAAFAR M A, GUEYE B, CANTIN F, et al. Towards a two-tier internet coordinate system to mitigate the impact of triangle inequality violations[C]//Proc of IFIP-TC6 Networking. Singapore: [s.n.], 2008.

[10] MAO Y, SAUL L, SMITH J M. IDES: an internet distance estimation service for large network[J]. IEEE Journal on Selected Areas in Communications, 2006, 24(12): 2273-2284.

[11] LUA E K, GRIFFIN T, PIAS M, et al. On the accuracy of embeddings for internet coordinate systems[C]//Proc of ACM IMC'2005. Berkeley, CA, USA: ACM, 2005.

编辑 黄莘