

增强虚拟网络坐标系统准确性的时延预处理技术^①

李岩岩, 张瑞林, 赖思恩, 吴桐树

(浙江理工大学 信息学院, 杭州 310018)

摘要: 在实际网络中网络时延不断抖动甚至出现长拖尾现象, 不仅会导致节点坐标振荡, 还严重影响虚拟网络坐标系统的稳定性及时延预测的准确性, 因此必须对实测时延进行预处理以避免“失真时延”影响虚拟网络坐标系统的性能. 介绍并分析了已有的时延预处理算法 Δ -filter 和 MP-filter 的性能特点, 提出了 RPE-filter 时延过滤算法, 实验对比表明 RPE-filter 算法不仅能够过滤遭受随机延迟污染的时延, 还能平滑时延曲线的波动, 更能实时反映网络拓扑的变化, 且绝对误差累加和明显小于 Δ -filter 和 MP-filter.

关键词: 网络坐标; 长拖尾; 时延过滤; 递归点估计; 集中域

Latency Pretreatment Technology to Improve the Accuracy of IP Network Coordinate Systems

LI Yan-Yan, ZHANG Rui-Lin, LAI Si-En, WU Tong-Shu

(School of Information Science and Technology, Zhejiang Sci-Tech University, Hangzhou 310018, China)

Abstract: In the actual network, network delay jitter even with a long tail, not only causes coordinates oscillation, but also seriously affects the stability of IP network coordinate system and accuracy of time delay prediction, so it is necessary to preprocess the measured time delay to avoid the performance of IP network coordinate system affected by “distortion delay”. This paper introduces and analyses the characteristics of the existing delay preprocessing algorithms of Δ -filter and MP-filter, then puts forward another delay filter algorithm called RPE-filter. Comparing experiment shows that RPE-filter can not only filter the time delay polluted by random delay, but also can smooth fluctuation of the delay curve, even more can real-time reflect the changes of network topology, and the most important is that its absolute error accumulation is obviously less than Δ -filter and MP-filter.

Key words: network coordinates; long tail; delay filter; recursive point estimation; concentrated domain

基于基准节点的 GNP^[1]、基于弹簧系统的 Vivaldi^[2] 等间接时延预测技术, 均采用空间嵌入原理, 将时延空间嵌入到几何空间, 构建虚拟网络坐标系统为诸如分布式数据库查询优化^[3]、覆盖层路由^[4]、资源最优放置^[5]等网络应用性能优化提供了新的途径. 虚拟网络坐标系统的构建以及节点坐标的更新都需要部分节点间的实测时延作为输入时延. 由于互联网络的复杂性、网络流量的动态变化和网络路由的动态选择, 网络时延随时都在不停的变化(称为抖动). 当网络出现拥塞、排队时间拖长以及拓扑变化时, 常常导致所测得的时延值大幅波动甚至跳跃, 如果直接采用“失真”的时延值作为输入时延, 那么节点定位的坐标位置被

错位, 系统时延预测性能不佳, 因而从时延样本中过滤出正确的时延值作为输入时延对提高时延预测性能至关重要.

采用正确的时延作为输入时延是确保虚拟网络坐标系统性能的前提, 本文总结了并分析了已有的时延过滤算法, 并提出了递归点估计时延过滤算法(RPE-filter). 经实验对比, RPE-filter 更能实时地、正确地反映网络的性能.

1 时延分析与系统流程图

PlanetLab^[6]从2005年5月2日开始用了超过3天的时间, 依托分布于北美重点大学和研究实验室的

^① 收稿时间:2012-09-13;收到修改稿时间:2012-10-23

269 个站点, 共包括 500 台机器, 通过每秒执行一次 traceroute 和 ping 命令获得一份共包含 43000000 个时延样本的数据集. 图 1 对此数据集中两节点间的时延样本进行累加概率分布统计, 其中几百毫秒的时延值占整个时延空间的比例较大, 但也有少量的时延值超过 1500ms, 大于通常情况下洲际链的时间延迟, 说明有长拖尾情况出现. 最小值与最大值之间的波动较大, 说明在实际的网络环境下出现时延值远远大于我们所期望的正常值. 如果直接采用实测时延样本作为虚拟网络坐标定位系统的输入时延, 系统将随着输入时延的波动而难以收敛, 同时时延预测的准确性很低, 因此需要对实测时延进行预处理, 从中过滤出合理的时延值作为定位系统的输入时延, 如图 2 所示.

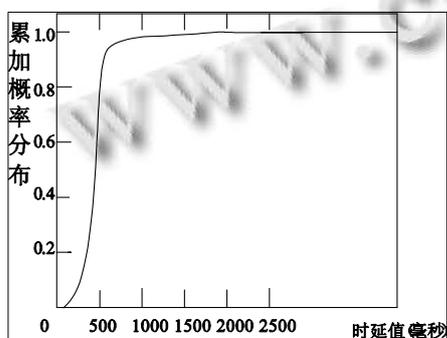


图 1 两节点间的时延累加概率分布^[7]

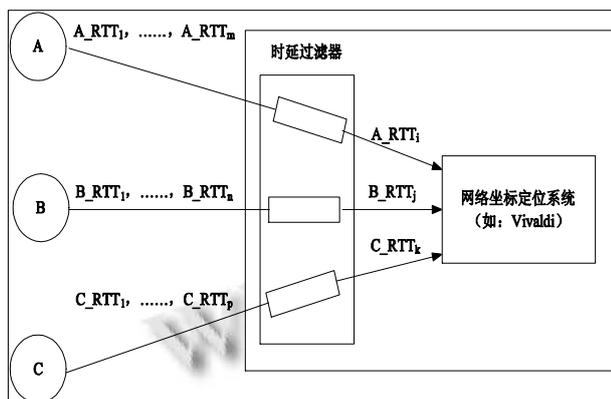


图 2 系统流程图

2 输入时延预处理方法

2.1 已有的时延预处理方法及其缺陷

文献[7]首次关注到了采用异常时延数据对网络坐标系统的稳定性和准确性的负面影响, 提出了基于滑动窗口机制的 MP-filter 时延过滤方法. MP-filter 是目前被广泛采用的抑制随机延迟污染的方法, 它有两个

参数: 一个是窗口的大小 W , 另一个是百分位 P . 首先将最近采集到的 W 个时延数据按照从小到大的顺利排列到窗口上, 然后选择第 $N(N=W \times P)$ 位的时延值作为输入时延, 通过实验表明 $P=25$ 的时延预测性能优于 $P=50$ 的时延预测性能. MP-filter 能较好地抑制时延的随机延迟污染, 但这种时延预处理方法没有维持原始时延样本总体的统计特征, 例如当窗口大小 W 为 4 且 $P=25$ 时, 即选择最小的时延作为坐标系统的输入时延, 过滤出的时延与样本间的偏差较大, 缩小了网络中节点间的距离, 所构建的系统预测时延值偏小.

文献[8]借鉴图像处理中的高斯平滑滤波^[9]的思想提出了 Δ -filter 时延过滤算法. 它由实测时延与其窗口权值进行迭代卷积求得坐标系统的输入时延, 其中窗口的权值取决于时延梯度并与时延梯度的 n 次方成反比关系, 时延波动幅度较大, 则相应的窗口权值自适应地减小, 达到平滑时延的目的, 抑制时延的波动. 窗口的权值函数决定了平滑的程度, 权值越小, 平滑作用越大, 但存在着边缘模糊的负面影响; 权值越大, 作用越小, 抑制污染的能力越差, 但边缘的定位越精确.

2.2 递归点估计时延过滤算法——RPE-filter

连续从山西省太原市古交市一节点 ping 节点 218.58.126.222 得到 584 个时延数据, 画出时延随时间变化的曲线如图 3.

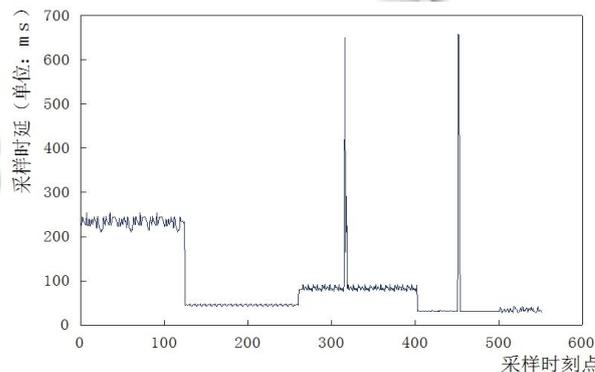


图 3 时延变化曲线

由图 3 可知, 节点间的时延是随着时间不断变化的, 包括三种变化趋势:

- (1) 单点随机跳跃, 通常远比正常值大得多, 如第 316 次采样得到的时延值;
- (2) 围绕一条水平线上下波动, 如从第 1 次到第 124 次采样得到的时延值;
- (3) 阶梯性变化, 包括从低台阶变化到高台阶的

“上升沿”和从高台阶变化到低台阶的“下降沿”两种情况, 分别如第 260 次和第 124 次采样前后的变化, 这可能是由于拓扑的变化引起的。

利用部分实测时延构建虚拟网络坐标系统提供时延预测功能是实时的服务, 这就要求输入时延应当能及时地正确地反映网络性能, 因此过滤实测时延要达到以下三个目的:

- 1) 能够去除随机跳跃的单点;
- 2) 定位波动中心并平滑波动曲线;
- 3) 保留阶梯性变化的边缘, 以实时地感应网络拓扑变化。

明确了时延预处理的目的, 本文基于概率论与统计学中的点估计^[10]的思想提出了递归点估计时延过滤算法。虚拟网络坐标系统每隔时间 T 更新一次节点的坐标, 假设在当前时间间隔 T 内按时间先后顺序测得的时延为 d_1, d_2, \dots, d_n , 由此构成一个具有先后顺序的序列集合 $E_1=\{d_1, d_2, \dots, d_n\}$, 其中 n 不是固定值。递归点估计时延过滤算法的过程如下:

按公式(1)计算时延样本均值 \bar{d} , 在无单点随机跳跃及阶梯性跳变的情况下, 样本均值是样本的重心, 即平衡点, 是总体均值很好的估计。依次比较 $d_i(1 \leq i \leq n)$ 与 \bar{d} 的大小, 仅出现一个大于是由单点随机跳跃引起的、有大有小且无规律则是由曲线的波动性引起的、先小后大是由上升沿、先大后小是由下降沿引起的。

按公式(2)计算时延平均偏差平方和 s_n^2 , s_n^2 表征了每个时间间隔内时延样本的分散程度, 并反映了时延样本的波动幅度。

$$\bar{d} = \frac{1}{n} \sum_{i=1}^n d_i \tag{1}$$

$$s_n^2 = \frac{1}{n} \times \sum_{i=1}^n (d_i - \bar{d})^2 = \frac{1}{n} \times \left(\sum_{i=1}^n d_i^2 - n \times \bar{d}^2 \right) \tag{2}$$

$$d_l = \bar{d} - s_n \tag{3}$$

$$d_h = \bar{d} + s_n \tag{4}$$

利用公式(3)和(4)求得样本的集中域, 即区间 E , $E=[d_l, d_h]$ 。将序列集合 E_1 中落在区间 E 中的时延数据按照原始时间先后顺利构成一个新的序列集合 $E_2=\{d_i, d_j, \dots, d_k\}$, 其中 $1 \leq i < j < k \leq n$, 此时的序列集合 E_2 与原始序列集合 E_1 有相同和不同两种情况。如果 E_2 与 E_1 相同, 且 $d_i(1 \leq i \leq n)$ 与 \bar{d} 的比较结果是曲线的波动性, 则直接输出作为输入时延, 算法结束; 如果 E_2

与 E_1 相同, 且 $d_i(1 \leq i \leq n)$ 与 \bar{d} 的比较结果是先小后大或者先大后小, 则将后半部分赋值给 E_2 , 此时 E_2 与 E_1 不同; 如果 E_2 与 E_1 不同, 则将 E_2 赋值给 E_1 , 递归调用上述过程。

RPE-filter 算法中并不直接输出当前时间间隔内测得的 n 个时延值的均值作为输出时延, 这是因为如果所测时延值中有跳跃值或者发生了阶梯性变化, 均值并不是正常情况下的时延值的无偏估计。因此将 n 个时延值与其均值作比较, 从中选择出更能反映当前网络状态的时延集合, 再递归调用上述过程, 从合适的时延序列集合中过滤出时延值的核心集中域集合, 直至时延集合不再变化, 最终集合的均值即为当前时延值的无偏估计, 作为系统的输入时延值。

2.3 各种时延预处理算法性能对比

分别采用 MP-filter、 Δ -filter、RPE-filter 这三种时延过滤算法对实测时延进行过滤, 其中设定 MP-filter 的窗口大小 $W=8$ 、百分位 $P=25$, Δ -filter 的窗口大小为 8、窗口权值与时延梯度的 2 次方成反比、迭代次数为 1, RPE-filter 的原始序列集合的长度为 8, 即每 8 个实测时延按照各自的算法输出一个时延作为虚拟网络坐标系统的输入时延, 为能在图中更加清晰可见, 这 8 个实测时延的一个输出时延采用相同的 8 个值表示, 如图 4 所示。

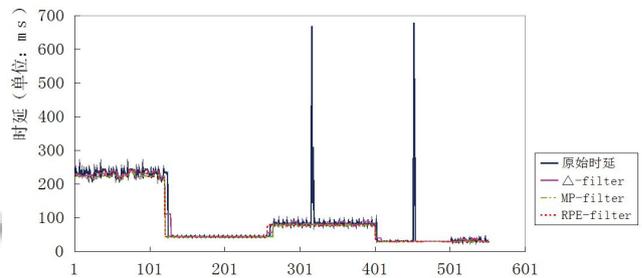


图 4 三时延过滤算法过滤结果对比

从图 4 可得出: Δ -filter 算法能够抑制随机单点跳跃, 也达到了平滑曲线波动性的目的, 但对时延阶梯性变化的感知性差, 不能及时反映网络拓扑的变化, 如在第 124 次时延采样所在的时间间隔内无法感知下降沿、在第 260 次采样前后的上升沿输出的值偏小; MP-filter 算法相当于输出在窗口中次小的值, 因此能够抑制随机单点跳跃, 也能够平滑曲线的波动性, 对如第 124 次采样前后的下降沿这样的阶梯性变化处理效果也较好, 但对如第 260 次采样前后的上升沿感知性最差, 用次小值平滑曲线却不考虑总体的统计特征,

输出值不是总体的无偏估计,使总体重心偏小;RPE-filter 算法能够准确剔除随机单点跳跃,用正常时延序列的重心平滑波动的曲线,且能及时感知因网线拓扑的改变而引起的时延阶梯性变化。

对同一个实测时延数据集计算每个时延过滤算法输出的时延值与当前实测时延的绝对误差,按照时间的先后顺序对绝对误差作累加和,得到如图 5 中的三种时延过滤算法的性能对比图,从此图可以看出 Δ -filter 的性能最差、MP-filter 的性能处于两者之间、RPE-filter 的性能最好。

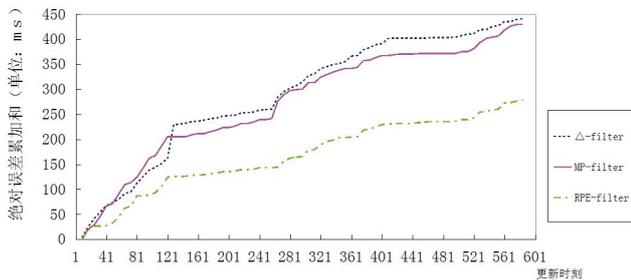


图 5 三种时延预处理算法性能对比

3 结语

从理论以及对时延数据集的统计分析得出在实际网络中网络时延出现抖动以及长拖尾现象,因此必须对时延样本进行预处理以避免“失真时延”对构建和维护虚拟网络坐标系统的性能产生负面影响.对已有的时延预处理算法的优缺点进行了分析,提出了递归点估计时延过滤算法,选择核心集中域的重心作为系统的输入时延.将本文算法与已有的两种时延过滤算法的性能进行了对比,本文算法不仅能够去除随机延迟

污染,还能平滑时延的波动,更能实时反映网络拓扑的变化,且绝对误差累加和明显小于已有的算法。

参考文献

- 1 T.Ng, Zhang H. Predicting internet network distance with coordinates-based approaches. Proc. IEEE INFOCOM, 2002: 170-179.
- 2 Dabek F, Cox R, Kaashoek F, et al. Vivaldi: A Decentralized Network Coordinate System. Proc. of ACM SIGCOMM'04. Portland: ACM Press, 2004. 15-26.
- 3 Pietzuch P, Ledlie J, et al. Network-Aware Operator Placement for Stream-Processing Systems. ICDE, 2006, Apr.
- 4 Andersen DG, Balakrishnan H, Kaashoek MF, Morris R. Resilient overlay networks. Proc. 18th ACM SOSP, 2001.
- 5 Pietzuch P, et al. Network-Aware Overlays with Network Coordinates. IWDDS, July 2006.
- 6 PlanetLab. <https://www.planet-lab.org/>, 2007-1-1/2012-10-8.
- 7 Ledlie J, Pietzuch P, Seltzer M. Stable and Accurate Network Coordinates. Proc. of International Conference on Distributed Computing Systems. Portugal, Computer Society. 2006. 84-93.
- 8 张敏,周亮,黄琼,阳小龙,隆克平.随机延迟污染对 IP 网络坐标系统的影响及其抑制方法研究.通信学报,2011,32 (12): 29-35.
- 9 李朝辉.数字图像处理及应用.北京:机械工业出版社,2004. 136-165.
- 10 茆诗松,贺思辉.概论论与统计学.武汉:武汉大学出版社, 2010. 235-271.

(上接第 94 页)

本文提出了基于 K-Means 聚类的人工鱼群算法,通过人工鱼群算法的全局搜索能力较强,且不易陷入局部最优的特点,改进了 K-Means 聚类对初始点选择不稳定的算法特性,大大的提高了全局寻优的能力,同时也具备了收敛速度快,效果好,且简单便捷的特点.仿真实验结果表明该算法的有效性与可行性。

参考文献

- 1 李晓磊,路飞,田国会,钱积新.组合优化问题的人工鱼群算法应用.山东大学学报(工学版),2004,34(5):64-67.
- 2 王闯.人工鱼群算法的分析及改进[硕士学位论文].大连:

大连海事大学,2008.

- 3 黄宇.基于 MARKOV 随机场 k-means 图像分割算法.电子学报,2009,37(12):2700-2704.
- 4 冯晓蒲,张铁峰.四种聚类方法之比较.微型机与应用, 2010,29(16):1-3.
- 5 李晓磊.一种新型的智能优化方法—人工鱼群算法.杭州:浙江大学,2003.
- 6 李晓磊,邵之江,钱积新.一种基于动物自治体的寻优模式:鱼群算法.系统工程理论与实践,2002,(11):32-38.
- 7 刘白,周永权.一种基于人工鱼群的混合聚类算法.计算机工程与应用,2008,44(18):136-138.