

基于坐标的网络节点聚类在 Internet 中的实验研究

陈 阳, 邓北星, 李 星

(清华大学 电子工程系, 北京 100084)

摘要: 在分布式网络应用中,节点聚类是构建高效网络体系结构的有效办法。网络坐标能够反映节点在 Internet 中的位置,利用网络坐标可以对网络节点之间的距离进行预测。在测量各个节点的网络坐标的基础上,利用 Triangulated heuristic 方法对节点间的网络距离进行预测,提出了网络距离作为参数对网络节点进行分布式聚类的算法,并在 PlanetLab 分布式实验床遍布全球的 156 个节点上进行了实际的测量和分析。分析结果表明,算法能够在网络规模较大的情况下对网络节点进行分布式聚类,算法有较高的可靠性和可扩展性。

关键词: 网络坐标; 网络距离预测; 节点聚类
中图分类号: TP393 **文献标识码:** A

0 引言

近年来随着 Internet 在全球范围的飞速发展,应用层组播、Peer to Peer 网络、大规模文件存取、文件多点下载等大规模分布式的网络应用在 Internet 服务中占的比重越来越大。在这些网络应用中,如何根据网络节点的分布状况构建一个高效的网络体系结构是非常重要的。利用网络节点之间的距离对网络节点进行聚类是一种行之有效的办法。

本文在测量各个节点的网络坐标的基础上,利用 Triangulated heuristic 方法对节点间的网络距离进行估计。提出了用估计得到的网络距离作为参数对网络节点进行聚类的算法。通过在 PlanetLab 分布式实验床^[1]遍布全球的 156 个节点上的实际测试,对算法在实际网络中的运行效果进行了分析。

1 网络坐标与网络距离

测量网络距离的最直接做法就是利用 ping 或者 traceroute 等工具进行端到端的测量。这种方法尽管实现简单,但是由于在网络规模比较大时会网络中各个节点增加比较大的测量负载,不具备可扩展性,不适合在大规模分布式网络中应用。

标识 Internet 节点位置的一种简单方法是通过测量该节点与 Internet 中一组预先选定坐标参考节点

(landmark) 的距离,这组值就作为这个节点的网络坐标。在 Internet 中的在线节点中,如果能够回应 ICMP Ping 的消息,并且本身比较稳定,均可选为坐标节点。

在测量得到坐标之后,就可以利用测得的坐标对网络节点之间的距离进行估算。目前基于网络坐标的距离预测大致有以下几种方法: IDMaps^[2]、Triangulated heuristic^[3]、GNP^[3]和 Vivaldi^[4]。本文采用 Triangulated heuristic 算法来估算距离。

Triangulated heuristic 是一种在假定 Internet 满足最短路径路由的前提下对网络距离进行预测的方法。其核心思想是在 Internet 上选取 N 个节点作为坐标参考节点 B_i 。对一个特定的节点 H ,通过测量 H 和 N 个坐标参考节点之间的距离可以得到节点 H 的坐标,为 $(d_{HB1}, d_{HB2}, \dots, d_{HBN})$ 的形式。对于需要估计相互之间网络距离的两个节点 H_1 和 H_2 ,近似认为 Internet 满足最短路径路由,所以三角不等式是成立的。Triangulated heuristic 规定 H_1 与 H_2 的距离下界 L 为

$$L = \max_i \{1,2, \dots, N\} (|d_{H_1B_i} - d_{H_2B_i}|) \tag{1}$$

H_1 与 H_2 的距离上界 U 为

$$U = \min_i \{1,2, \dots, N\} (|d_{H_1B_i} + d_{H_2B_i}|) \tag{2}$$

于是 L 和 U 的不同加权平均就能用来作为距离函数来预测 H_1 与 H_2 之间的距离。

收稿日期: 2005-12-07.
基金项目: 国家自然科学基金资助项目(60473087).
作者简介: 陈 阳(1981-), 男, 博士生, E-mail: chen yang04@mails.tsinghua.edu.cn; 邓北星(1964-), 男, 博士, 副教授, E-mail: dengbx@mail.tsinghua.edu.cn; 李 星(1956-), 男, 博士, 教授, 博士生导师, CERNET 网络中心副主任, E-mail: xing@cernet.edu.cn.

作者利用相对误差来评价网络距离的预测效果:

$$\text{相对误差} = \frac{|\text{预测距离} - \text{实际距离}|}{\min(\text{预测距离}, \text{实际距离})} \quad (3)$$

预测距离等于实际距离时相对误差为 0, 相对误差越接近 0 表明预测的效果越好.

Guyton 和 Schwartz 为了解决最近服务器选择问题, 用 $(L + U)/2$ 作为距离估计函数, 仅取得了有限的成功^[5]. Ng 和 Zhang 于 2001 年 5 月在全球范围内随机抽取了 869 个节点作为测量对象, 取了 15 个坐标参考节点, 对 Triangulated heuristic 方法的可靠性作了非常细致的研究. Ng 和 Zhang 的研究结果表明, 上限 U 实际上有很好的精确性, 采用上限 U 作为估计函数, 在该方法预测的所有路径中, 90% 的相对误差小于 0.59, 其效果比下限 L 或者上下限平均 $(L + U)/2$ 都好很多^[3]. 因此, 本文采用上限 U 作为估计 H_1 和 H_2 之间距离的函数.

2 基于坐标的节点聚类

通过以上的办法得到网络节点之间的距离之后, 可以对网络节点进行聚类.

本文聚类算法的设计基于如下的假设: 网络节点在网络中呈聚类分布, 聚类内部距离小于聚类之间的距离, 聚类之间的链路带宽远比聚类内部稀缺. 根据当前的 Internet 结构, 这一假设是普遍存在的^[6]. 因此, 在很多大规模分布式网络应用中减少数据在聚类网络之间的传送, 可以提高整体的传输效率. 为了构建高效的网络体系结构, 需要根据网络距离对网络节点进行聚类.

在上述的聚类算法中, 每一个聚类网络由网络距离相近的网络节点形成. 相隔很远的节点分别加入不同的聚类网络. 每一个聚类网络中有一个节点担任该聚类的根节点. 所有的根节点形成上层的核心网络.

考虑到实际网络中节点可能随时加入或退出, 因此需要采用分布式的聚类方法. 节点的分布式聚类过程如下.

当一个新的节点开始加入的时候, 它首先在核心网络中查询具有坐标向量与自己的坐标向量利用 Triangulated heuristic 方法估算出来网络距离最小的根节点, 作为离自己最近的根节点. 如果这个新节点到离自己最近的根节点的网络距离大于一个设定的阈值 (称为聚类半径), 说明该参与节点到达任何一个聚类网络的距离都很远, 因此它加入核心网络并作为新的聚类网络根节点. 如果到最近根节点的距离小于聚类半径, 则该节点加入最近根节点所在的聚类网络.

此处需要说明两点: (1) 聚类半径由节点的分布和网络状况事先设定. 在实际的环境中, 如果需要将相同 ISP 节点聚在同一个聚类网络中, 则聚类半径就是各个 ISP 的平均半径; (2) 存在个别节点, 由于其网络坐标每一维的值都大于聚类半径, 在本文的基于上限 U 的 Triangulated heuristic 方法中, 是不可能和其他任何一个节点加入同一个聚类的. 这部分节点作者定义为“孤岛”. 它们将不会加入核心网络, 也不会加入到任何一个聚类当中.

在节点聚类过程中, 存在的主要误差包括: (1) 节点的坐标向量不能完全精确标识节点在网络中的位置. 相互之间实际网络距离很小的节点, 对应的坐标向量可能差异很大. 为了减少此类误差, 需要增加坐标参考节点的数目, 并且应选择相互之间相隔尽量远的节点作为坐标参考节点, 从而提高坐标向量的准确性. (2) 基于上限 U 的 Triangulated heuristic 方法本身存在的误差, 因为 Triangulated heuristic 方法的前提是 Internet 满足三角形法则, 这一点在实际的 Internet 中由于低效率路由的存在, 并不是总能成立的. Francis 等的研究表明, 对于所有要测量的三角闭合回路 (a, b) 、 (b, c) 和 (a, c) , 计算的 $\frac{(a, c)}{(a, b) + (b, c)}$ 的比值, 有 7% 的比值比 1 大^[7], 这与其他的一些研究是一致的. 这一误差也会对网络距离预测的准确性造成影响.

基于坐标的节点聚类算法在实际网络中的实现效果, 下面将通过在 PlanetLab 上进行的实验进行分析.

3 PlanetLab 的实验结果

本研究的实验是基于 PlanetLab 进行的. PlanetLab 是一个全球性的拥有 200 多个节点、近 600 台互联的计算机的分布式实验床, 参与节点遍布全球 100 多所著名大学和科研机构, CERNET 有 25 所高校于 2004 年 12 月加入了 PlanetLab. 利用 PlanetLab 可以进行互联网的一系列实际的大规模的测试.

作者于 2005 年 6 月, 在遍布全球的 156 个 PlanetLab 节点上对提出的基于坐标的分布式网络节点聚类算法进行了实际的测试. 坐标测量的参考节点有 14 个; 其中亚洲 5 个, 北美 5 个, 欧洲 4 个. 作者选取的参考聚类半径为 50 ms.

分布式聚类算法的运行结果如下:

156 个节点中, 有 4 个节点属于“孤岛”; 其他节点分布在 5 个聚类中, 每个聚类的核心节点的 IP 地址列入表 1.

表 1 在 PlanetLab 上的聚类结果

Tab. 1 Clusters in PlanetLab

编号	聚类根节点	节点个数
1	128.31.1.11(MIT)	53
2	128.111.52.61(UCSB)	38
3	128.232.103.201 (Cambridge)	25
4	140.112.107.80 (Taiwan)	6
5	202.38.99.68 (CERNET)	30

从表 1 的结果可以看到,所有的 156 个节点除了 4 个孤岛(占 2.56%)之外,其他节点所处的聚类

和节点的实际分布情况基本上是相符的.对参与聚类的 152 个节点进行分析,聚类 1 基本上都是美国东海岸和加拿大的节点,聚类 2 基本上是美国西海岸包括中部的一些节点,聚类 3 都是欧洲的节点,聚类 4 都是台湾地区和亚洲其他国家的节点,聚类 5 都是 CERNET 的节点.在初步分析的基础上,作者利用定量的实际测量来检验聚类算法的效果.

根据聚类的结果,让各个节点对于网络中其他节点发起测量,测量得到各个聚类之间的平均距离(对角线上的元素表示同一个聚类内部节点之间的平均距离)如表 2.表 2 中,对角线上的元素大小远远小于表中其他元素的大小,说明同一聚类节点之间的距离远远小于不同聚类节点之间的距离,因而聚类的效果是比较可靠的.

表 2 各个聚类之间的平均距离

Tab. 2 Average distances between Clusters

	1	2	3	4	5
1	31	79	107	202	249
2	79	20	162	157	225
3	107	162	31	292	341
4	202	156	291	4	75
5	249	225	341	75	30

由于本文的聚类采用的是分布式聚类算法,而

且并不涉及节点之间的端到端测量,不会在网络规模比较大时对网络中各个节点增加比较大的测量负载.所以,算法本身具有较高的可扩展性.

4 结 论

通过在 PlanetLab 分布式实验床上进行的较大规模的实际网络测量,可以得出以下结论:本文提出的利用 Triangulated heuristic 方法进行网络坐标和距离的计算,并在此基础上对网络节点进行聚类的算法,能够在网络规模较大的情况下对网络节点进行分布式聚类.算法有较高的可靠性和可扩展性.

参考文献：

[1] PlanetLab Website [EB/OL]. [S. l.]:[s. n.], 2005 [2005-08-05]. <http://www.planet-lab.org/>.

[2] PAUL F, SUGIH J, CHENG J, et al. IDMaps: A global internet host distance estimation service[J]. *IEEE/ACM Transactions on Networking*, 2001, 9(5):525-540.

[3] NG T S E, ZHANG H. Predicting Internet network distance with coordinates-based approaches [J]. *Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies*, 2002, 1:170-179.

[4] FRANK D, RUSS C, FRANS K, et al. Vivaldi: A decentralized network coordinate system[A]. *Proceedings of ACM SIGCOMM 2004*[C]. Portland:[s. n.], 2004.

[5] GUYTON J D, SCHWARTZ M F. Locating nearby copies of replicated Internet servers [A]. *Proceedings of ACM SIGCOMM 1995*[C]. Cambridge:[s. n.], 1995.

[6] 田瑞雄. 自组织覆盖网络建模与优化[D]. 北京:清华大学, 2005.

[7] FRANCIS P, JAMIN S, PAXSON V, et al. An architecture for a global Internet host distance estimation service[J]. *Eighteenth Annual Joint Conference of the IEEE Computer and Communications Societies*, 1999, 1:210-217.

Experimental study on network coordinate-based node clustering

CHEN Yang, DENG Bei-xing, LI Xing

(Department of Electronic Engineering, Tsinghua University, Beijing 100084, China)

Abstract : In the application of large-scale distributed network, node clustering is a useful way to construct an effective network infrastructure. Network coordinate can reflect node's position in Internet and can be used to predict the network distance. After getting every node's network coordinate, the authors use Triangulated heuristic to predict the network distance between nodes. The authors present a distributed clustering algorithm using network distance, and having done the experiments on 156 nodes of PlanetLab testbed all over the world. The experimental results show that this algorithm can work well on large-scale node clustering in distributed way, and it has high reliability and scalability.

Key words : network coordinate; network distance prediction; node clustering