

基于 Hilbert 曲线的拓扑匹配的 P2P 覆盖网模型

李 永^{1,2}, 余镇危²

(1 盐城师范学院 信息科学与技术学院, 江苏 盐城 224002; 2 中国矿业大学(北京) 机电与信息工程学院, 北京 100083)

摘 要: 针对 P2P 网络中由于逻辑网络和物理网络的拓扑结构不匹配导致物理路由效率低下的问题, 提出一种新的拓扑匹配的 P2P 覆盖网模型. 首先基于 Vivaldi 网络坐标系统对网络节点进行聚类, 划分成 K 个聚集, 且在每个聚集内选出头节点; 然后利用 Hilbert 空间填充曲线的局部保持特性, 把 K 个聚集的头节点构成环状拓扑结构; 最终得到一个拓扑匹配的 Hilbert-Ring 覆盖网模型. 仿真实验表明, 该模型具有良好的性能, 可以有效地降低网络延迟, 减少网络开销.

关键词: 覆盖网; Hilbert 曲线; 拓扑匹配

中图分类号: TP393

文献标识码: A

文章编号: 1000-7180(2013)01-0139-04

A Topology-Matching P2P Overlay Network Model Based on Hilbert Curve

LI Yong^{1,2}, YU Zhen-wei²

(1 School of Information Science and Technology, Yancheng Teachers University, Yancheng 224002, China;

2 College of Mechanical and Electronic Engineering, China University of
Mining and Technology (Beijing), Beijing 100083, China)

Abstract: Aiming at the unmatched topology problem between overlay network and physical network which result in inefficient routing, a new topology-matching P2P overlay network model is presented. Firstly, network nodes are clustered into K groups based on the network coordinates system Vivaldi, and selecting a leader node for each group. Then, using the proximity-aware nature of hilbert space filing curve, the K leader nodes are formed into ring topology. Finally a topology-matching Hilbert-Ring overlay network model is obtained. Simulating experiment shows that this model has the good performance, and it is effective in lowering the network delay and decreasing the network load.

Key words: overlay network; hilbert currve; topology-matching

1 引言

P2P 网络具有高可扩展性、容错性、高性能, 以及能充分利用网络边缘端系统节点资源等特点, 在 Internet 应用中备受关注. 近年来, 对覆盖网(Overlay)的研究已经成为 P2P 领域的研究热点之一. 出现了许多基于覆盖网的应用, 如 Napster, Gnutella, OceanStore, Kademlia 等文件共享系统; Split-Stream, PROMISE, Bullet 等大型数据分发系统; 基于分布式哈希表(DHT)的结构化 P2P 应用以及内

容分发网络(CDN)等^[1]. 在 P2P 系统中, 覆盖网络建立在逻辑概念上, 逻辑上最近的路径并不一定在物理上最近, 因此, 节点不能与其物理相邻近的节点交换数据, 造成上层应用的数据传送和资源定位低效, 同时也对下层物理网络造成巨大的带宽压力. 这种逻辑层与物理层上的结构差异问题被称为拓扑失配问题. 针对该问题, 文献[2]利用每个节点周期性地向邻居获取它们邻居拓扑信息, 选择部分距离最近的节点作为自己新的邻居. 但在网络规模大时, 存在控制消息代价高、收敛速度慢等问题, 且不能保证

结果为全局最优. SAT-Match^[3]是一种针对结构化覆盖网络 CAN 的局部优化措施,包括探测阶段和跳转阶段.由于结构化覆盖网络中节点状态与节点位置密切相关,使得 SAT-Match 优化过程需要大量跳转操作、状态迁移和探测报文,系统代理负荷沉重.文献[4]直接从 BGP 路由器上获取拓扑信息,节点根据它们的 IP 网络地址被分成簇.但由于 Internet 网络中端节点没有 BGP 路由表的访问权限,该方法对普通端节点根本无法适用.文献[5]通过节点的网络坐标并计算出与其他节点之间的网络距离,根据网络距离使用 K-means 算法进行网络节点聚类,但由于随机选取聚类数和初始聚类中心,对聚类结果影响较大.

针对 P2P 覆盖网的拓扑失配问题,本文首先对网络节点进行聚类,划分成 K 个聚集,且在每个聚集内选出头节点;然后利用 Hilbert 空间填充曲线的局部保持特性,把 K 个聚集的头节点构成环状拓扑结构;最终得到一个拓扑匹配的层次环状拓扑的覆盖网模型.模拟实验表明,该模型可以有效地降低网络延迟,减少网络流量.

2 Vivaldi 和 Hilbert 空间填充曲线

2.1 Vivaldi

网络测量的相关研究表明,可以将因特网拓扑建模为一个多维欧几里得空间.本文采用网络坐标系统 Vivaldi^[6]将网络节点映射到 m 维坐标空间中,Vivaldi 是一种简单的、轻量级的网络坐标系统,具有良好的扩展性.在一个具有 n 个节点的网络中,每个网络节点 i 都有一个来自 Vivaldi 的 m 维网络坐标 $(v_{i1}, v_{i2}, \dots, v_{im})$, $i = \{1, 2, \dots, n\}$,表示网络节点映射为 m 维网络坐标空间 R^m 中的一个点.设节点 j 的坐标为 $(v_{j1}, v_{j2}, \dots, v_{jm})$,则通常情况下,若 $(v_{i1}, v_{i2}, \dots, v_{im})$ 与 $(v_{j1}, v_{j2}, \dots, v_{jm})$ 在 m 维空间的距离较近,则节点 i 与节点 j 在网络中的实际距离也较近.

2.2 Hilbert 空间填充曲线

空间填充曲线是一条通过一个有限空间的所有点的曲线,它提供了一种多维欧几里得空间到 1 维欧几里得空间的映射方法.目前,这种映射关系被应用到许多的领域,包括图像的处理、多维数据的索引等.

定义 1 如果映射 $f: I \rightarrow E^d$ ($d \geq 2$) 为连续的,并且 $f(I)$ 是整个 d 维单位超立方体空间,那么称 $f(I)$ 为空间填充曲线,其中 E^d 为 d 维欧几里得空间,

$I = [0, 1]$.

空间填充曲线主要有 3 种: Z 序、格雷码和 Hilbert 空间填充曲线.研究表明,在这些空间填充曲线族中,HSFC(Hilbert Space Filing Curve)具有最好数据聚集特性,即 HSFC 能最好地保持空间位置的邻近特性,两个在 1 维空间相邻的点在 d 维空间必然相邻^[7].

对于 $m \geq 1, d \geq 2$,则 H_m^d 表示 d 维 HSFC 的 m 阶近似,它可以将整数区间 $[0, 2^{m \times d} - 1]$ 映射到 d 维整数空间 $[0, 2^m - 1]^d$.在实际应用中,一般只对有限粒度的 d 维空间感兴趣,因此只需构造 d 维 HSFC 的 m 阶近似即可.

定义 2 如果用 R^d 来表示 d 维欧氏空间,用 $x_1, x_2, \dots, x_d$, 来表示 d 维空间 R^d 的 d 个维,把从 d 维空间坐标转换成的 1 维空间坐标值叫做 Hilbert 编码,记为 H_{code} .

关于 Hilbert 曲线编码的生成有两种实现方法,一个是表驱动方法,表驱动方法通过扫描代码扫描列表来实现曲线生成.另一个是计算的方法.本文采用文献[8]给出的一种具有较好计算性能的方法.

3 拓扑匹配的覆盖网模型构造

3.1 网络节点的聚类

网络节点聚类问题可以用图论来描述,对一个加权完全图 $G = (V, E, W)$,其中 V 为顶点的集合, E 为边集, $W_G(u, v)$ ($u, v \in V$) 为连接顶点 u, v 的边上的权重,节点聚类问题就是求图 G 的尽可能少的 K 个子图 $G_1(V_1, E_1, W_1), G_2(V_2, E_2, W_2), \dots, G_K(V_K, E_K, W_K)$ 满足: (1) $V_i \cap V_j = \emptyset$ $i \neq j$; (2) $V_1 \cup V_2 \cup \dots \cup V_K = V$; (3) $W_i(u, v) < \epsilon, u, v \in V_i, i = 1, \dots, K$.

节点聚类问题的求解是困难的,而且实际应用中还需要测量节点间的两两距离,仅是节点间两两距离的测量所带来的消耗也是难以接受的,对此不少研究者提出了一些解决办法.在这里采用基于免疫克隆算法和 K-medoids 的网络节点聚类方法^[9],把 P2P 网络划分为 K 个聚集.

3.2 聚集内头节点的选举问题

聚集的头节点主要有以下三个功能: (1) 保存并维护自身成员信息及邻居聚集信息; (2) 接受新节点、聚集内成员或其它聚集的头节点等的信息查询; (3) 作为新加入节点的地理位置参照点.由于头节点作为整个聚集的代表,是新节点加入时的位置参照点,为了能使后来加入的节点加入到正确的聚集,

头节点应该始终是聚集中所有节点的中心,即应该是距离聚集中所有节点都比较近的节点.即头节点的选择依据为是否最靠近聚集所有节点分布的地理中心.因此,头节点的选择问题可归结为图论中完全赋权图的重心问题.

定义 3 设图 $G(V, E, M, L)$ 为完全赋权图.其中:顶点集 $V = \{v_1, v_2, \dots, v_n\}$;边集 $E = \{e_1, e_2, \dots, e_m\}$;顶点的权值集合 $M = \{m_1, m_2, \dots, m_n\}$,顶点 v_i 的权值为 m_i ,泛指质量;边的权值集 $L = \{l_1, l_2, \dots, l_m\}$,边 e_j 的权值为 l_j ,泛指边的长度.设 $d_{i,j}$ 为两顶点 v_i, v_j 间路径的长度(即两点间路径上边的权值和), $S_j(G) = \sum_{i=1}^n m_i \cdot d_{i,j}$ 定义为点 v_j 对应的功,则完全赋权图 G 的重心定义为: $P = \{v_k | v_k \in V, S_k(G) = \min_{v_j \in V} S_j(G)\}$.

如果将聚集中的节点看作图 G 的顶点,连接两个节点的底层网络路径作为图 G 的边,则图 G 是一个完全图.由于在头节点选择中,每个节点除了它的位置以外都看作是等同的,因此可以将它们的质量均定义为 1,如果再将两点间的距离作为对应边的长度,则图 G 就构成一个完全赋权图.聚集头节点的选择问题就转化为求完全赋权图 G 的重心问题.完全赋权图重心可采用文献[10]给出的矩阵验算法求得.该方法如下:令 $D = (d_{i,j})_{n \times n}$ 为完全赋权图 G 的最短距离矩阵, $M = \{m_1, m_2, \dots, m_n\}$ 为顶点的质量向量,则完全赋权图 G 的重心为向量 $W = M \cdot D$ 的最小元素对应的顶点.由于聚集中每个节点的质量均为 1,即 $M = \{1, 1, \dots, 1\}$,因此 $W = M \cdot D = (1, 1, \dots, 1) \cdot D = (\sum_{j=1}^n d_{j,1}, \sum_{j=1}^n d_{j,2}, \dots, \sum_{j=1}^n d_{j,n})$.可以看出,向量 W 的每个分量为对应节点到所有其它节点的距离之和.

3.3 基于 Hilbert 曲线的聚集间的覆盖网构造方法

由于 Hilbert 空间填充曲线的高聚集特性,使得 Hilbert 编码相近的节点其实际网络中也必定相邻,这样,通过节点的 Hilbert 编码就可以判断节点间的相对位置关系.不同节点所对应的 Hilbert 编码的差的大小,反映了这些节点之间的距离远近,因此每个节点只要保存自身的 Hilbert 编码,就等于保存了节点自身在网络中的相对位置信息.将每个头节点的 Vivaldi 坐标经 HFSC 映射产生的 H_{code} 作为聚集的 ID 号,然后将所有聚集按其 ID 的大小顺序排列构成一个环,如图 1 所示.即聚集间的覆盖

网拓扑结构是一个由聚集头结点构成的环,环上的每个聚集维护两条到其两个环上邻居聚集的链接,形成一个双向环,称为 Hilbert-Ring 模型.

为了维护聚集间覆盖网的环形拓扑结构,聚集的头节点要维护两个列表,一个是本聚集所有成员列表,一个是聚集邻居列表.成员列表包括聚集中所有成员的相关信息,其第一条记录为头节点自身的信息;邻居列表一般要包括邻居的 ID 及其头节点的 IP 地址等信息.一般情况下,邻居列表的第一条记录是 ID 比自己的 ID 小的环上邻居,称为左邻居,第二条记录记录 ID 比自己的 ID 大的环上邻居,称为右邻居.有两个特殊的聚集不符合上面的情况,一个是环上 ID 最大的聚集,另一个是环上 ID 最小的聚集,这时环上 ID 最小的聚集必然是 ID 最大聚集的右邻居,而 ID 最大的聚集必然是 ID 最小聚集的左邻居.当整个网络只有一个聚集时,其左右邻居都指向自己.

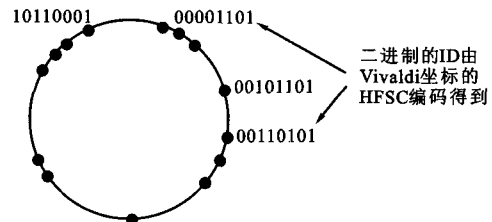


图 1 聚集间覆盖网的环形拓扑结构

4 仿真实验

本文评价 P2P 覆盖网与底层物理网络匹配性能的指标有:(1) 伸张度(Stretch).指从一个数据源节点到一个目的节点在覆盖网络的路径长度或者延时和在直接单播得路径长度或者延时的比例.这个参数可以用来形容 P2P 覆盖网逻辑拓扑与底层物理网络拓扑的相关度,stretch 越小说明与底层拓扑越相近.(2) 压力度(Stress).压力度是计算拓扑网络中经过一条链路或者路由节点相同数据包的次数.在实际数据传输的过程中,一个数据包如果多次被复制的话,则肯定会占用一部分带宽,一旦被复制的数据包急剧增多,网络带宽可能就会被耗尽.

利用 NS2 来仿真 Hilbert-Ring 模型构建方法,利用 GT-ITM 生成能代表当前 Internet 结构的 Transit-Stub 模型.实验中生成的底层物理网络由 10000 个节点组成,P2P 覆盖网络中的节点数分别选择 500、1000、1500、2000、2500、3000 进行实验.为了体现 Hilbert-Ring 模型的有效性和高效性,把

Hilbert-Ring 模型与聚集头节点间直接采用 Chord 协议的模型进行比较,在相同环境下 Hilbert-Ring 模型与 Chord 模型在以上两个指标上的曲线如图 2 和图 3 所示.图 2 所示的是随 P2P 覆盖网络节点数的不同的 Stretch 值的变化情况,从图上可以看出,Chord 模型具有不稳定性,而 Hilbert-Ring 模型的 Stretch 值较小且总体比较平稳.这证明 Hilbert-Ring 模型较 Chord 模型有很大的改进.图 3 描述的是随 P2P 覆盖网络节点的增多,一个数据包在网络被复制的次数. Hilbert-Ring 模型接近线性增长,而 Chord 模型当 P2P 覆盖网络节点数目增多的时候,复制数据包的次数急剧增多,造成网络开销很大.

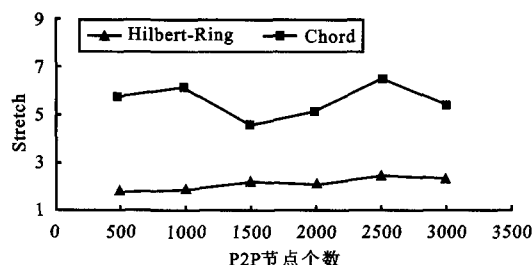


图 2 伸张度 Stretch 值比较

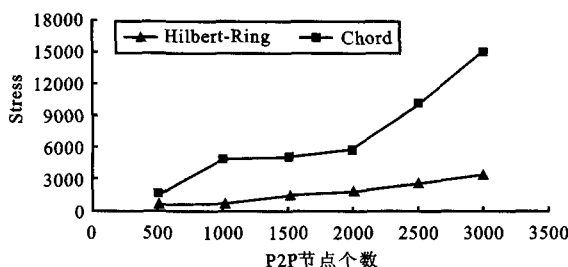


图 3 压力度 Stress 值比较

5 结束语

针对目前 P2P 覆盖网中存在的逻辑层与物理层拓扑不匹配的问题,构造了一种可以进行拓扑感知的 P2P 覆盖网模型——Hilbert-Ring 模型.其基本思想是根据底层物理网络的拓扑来建立上层的逻辑覆盖网.基于 Vivaldi 网络坐标系统对网络节点进行聚类,划分成 K 个聚集,且在每个聚集内选出头节点;为了减少聚集间的路由延时和开销,利用 Hilbert 空间填充曲线的局部保持特性把 K 个聚集

的头节点构成环状拓扑结构. Hilbert-Ring 模型的优势在于物理层与逻辑层的对应关系比较精确,可以有效地降低网络延迟,减少网络开销.

参考文献:

- [1] 王意洁,李小勇.网络距离预测技术研究[J].软件学报,2009,20(6):1574-1590.
- [2] Ceccanti A, Jesi G P. Building latency-aware overlay topologies with quick peer[C]//Proceedings of ICNS. USA:[s. n.],2005.
- [3] Ren S, Guo L, Zhang X. SAT-match: a self-adaptive topology matching method to achieve low lookup latency in structured P2P overlay networks[C]//Proceedings of the 18th International Parallel and Distributed Processing Symposium(IPDPS'04),2004.
- [4] Nakao A L, Bavier P A. A routing underlay for overlay networks [C]//Proceedings of IEEE SIGCOMM. Karlsruhe, Germany: IEEE,2003.
- [5] Jiang Yi, Shi Hongbo, You Jinyuan. Network coordinates-based peer-to-peer hosts clustering [C]//Proceedings of IFIP Internet Conference on Network and Parallel Computing. Tokyo, Japan: IEEE Press,2006.
- [6] Dabek F, Cox R, Kaashoek F, et al. Vivaldi: a decentralized network coordinate system [C]//Proc of ACM SIGCOMM'04. [s. l.]: ACM Press,2004.
- [7] Chen H, Chang Y. Neighbor-finding based on space-filling curves[J]. Information Systems, 2005, 30(3): 205-226.
- [8] 李晨阳,张扬,冯玉才. N 维 Hilbert 编码的计算[J]. 计算机辅助设计与图形学学报,2006,18(7):1032-1038.
- [9] 李永,余镇危. 基于免疫克隆算法和 K-medoids 的网络节点聚类[J]. 微电子学与计算机,2011,28(8): 119-122.
- [10] 杜端甫. 运筹图论(图、网络理论中的运筹问题)[M]. 北京:北京航空航天大学出版社,1990.

作者简介:

李永男,(1978—),博士研究生.研究方向为网络体系结构、覆盖网.

余镇危男,(1942—),教授,博士生导师.研究方向为网络体系结构、覆盖网、下一代互联网.