

基于节点服务能力的自适应 P2P 模型^①

陈水平 吴开贵 (重庆大学 计算机学院 重庆 400044)

摘要: P2P 网络的节点在处理能力、存储能力、网络带宽、兴趣域等特性上存在着明显的异构性, 在早期的研究中, 这种异构性往往被忽略或者仅仅简单、单一地被考虑。综合利用节点的这些异构性, 提出基于节点服务能力的自适应 P2P 模型。模型中, 节点通过服务能力自主选择邻居进行聚类, 根据网络的需求和动态变化自适应地调整聚类的规模和节点间的连接。仿真实验结果表明, 该模型减少了网络拓扑失配和 freerider 现象, 有效地提高了网络的搜索效率。

关键词: P2P; 异构性; 服务能力; 自适应; 模型

A Self-Adaptive Peer-to-Peer Model Based on Peer's Serve Capabilities

CHEN Shui-Ping, WU Kai-Gui

(College of Computer Science, Chongqing University, Chongqing 400044, China)

Abstract: In a peer-to-peer network, the peer's characteristics have significant heterogeneity on the processing power, storage capacity, network bandwidth, and the domain of interest etc. In the early studies, this heterogeneity was often overlooked, or just too simple and single to be considered. Synthetic utilization of these heterogeneous peers proposes a self-adaptive P2P model based on the peer's service capabilities. In this model, the peer chooses his/her neighbors and clusters through the service capabilities, accordingly with the network's needs of a dynamic self-adaptive adjustment of cluster size and the connections between peers. The simulation results show that the model reduces the network topology mismatch and free rider phenomenon, effectively improving the search efficiency of network.

Keywords: P2P (peer-to-peer); heterogeneity; serve capabilities; self-adaptive; model

1 引言

计算机对等网(Peer-to-Peer network, P2P)技术是目前计算机网络技术研究领域的一个热点。它具有自主性、容错性、高效性等好的特点, 并在分布式计算、文件共享、电子商务等领域都得到了广泛的应用。目前的非结构化 P2P 模型, 采用随机的节点组织方式, 节点在网络拓扑中的地位完全平等, 虽然这种结构具有组织策略简单、支持复杂查询等优点, 但是完全忽略了节点的异构性, 使得查询效率低、带宽消耗大, 不利于网络的扩展, 严重影响了 P2P 网络的可用性。

大量的研究表明 P2P 网络除具有明显的动态特性

外, 节点在处理能力、存储能力、网络带宽、兴趣域等特性上往往也存在较大的异构性。本文综合考虑节点的这些异构性, 提出基于节点服务能力的自适应 P2P 模型 BSC-P2P(Based on Serves Capabilities P2P)。其主要思想是(1)根据节点的异构性定义节点的服务能力, 节点按服务能力自组织进行聚类, 并从中选出服务能力最强的节点作为超级节点管理该组。(2)按网络实际需求动态地调整和维护节点之间的连接, 自适应网络的动态变化。(3)引入超级节点备份、超级节点代理冗余机制, 减少由于它们失败造成的网络部分瘫痪的概率。

^① 基金项目:国家自然科学基金(90818028)

收稿时间:2010-01-18;收到修改稿时间:2010-03-08

本文第 1 节是相关研究工作,第 2 节给出了模型的结构及相关概念,第 3 节具体描述了模型的核心算法,第 4 节用实验对模型进行了验证,最后总结全文并提出下一步工作。

2 相关工作

不同的策略用于 P2P 网络的组织。文献^[1]根据节点所共享内容的异构性来组织节点的物理拓扑,内容相似的节点相互连接组成一个语义网。这种网络仅仅考虑到了节点共享内容的异构性,没有考虑到节点其它的异构性,当语义组内的节点所要查找的内容与本组不相关的时候,必须使用泛洪的方法在整个网络中搜索所需要的资源,因此网络效率不高。文献^[2]、文献^[3]根据节点兴趣的异构性来构建 P2P 网络,相同兴趣的节点相互连接形成兴趣组以提高网络的搜索效率。但网络没有考虑节点物理位置上的异构性,容易造成拓扑失配,给网络通信带来严重的开销。文献^[4]在利用节点兴趣异构性构建 P2P 网络的基础上,引入中心节点备份机制,并在搜索算法中加入缓冲池(Cache)策略,有效的提高了网络的搜索效率和成功率,但是同样也没有考虑到节点物理位置及其它的异构性。

文献^[5-7]则从节点物理位置的异构性来组织 P2P 网络的拓扑。文献^[5]提出了一种基于节点物理位置信息的拓扑失配解决方法 LTM (Location-aware Topology Matching)。这种方法采用节点之间的通信延迟作为衡量节点位置远近的标准,根据节点间的位置远近,在保证搜索效率的前提下,P2P 网络中中长距离的低效连接被逐步剔除,代之以近距离的高效连接。文献^[6]将节点物理位置上的异构性和节点共享资源的异构性结合起来考虑,通过节点的物理位置来划分区域以减少物理网络和覆盖网络的拓扑失配问题,然后在区域内按节点共享资源的异构性来组织节点以提高网络的搜索效率。文献^[7]采用分布式哈希表机制将主题相同的节点组织成主题区域,在同一主题区域内根据物理位置相近原则进行群的划分,从而在相同主题的基础之上建立拓扑匹配的语义 P2P 覆盖网络。

以上这些方法仅仅是简单的利用到了节点某一、两方面的异构性,没有全面考虑到节点所有的异构性,而本文提出的基于节点服务能力的自适应 P2P 模型,

综合利用了节点多方面的异构性,使得网络在提高搜索效率、减少拓扑失配和 freerider 节点等方面都有了明显的改善。

3 BSC-P2P模型的结构

3.1 模型的基本概念

定义 1. 服务组。服务组是由超级节点和普通节点组成的自治系统,根据节点的服务能力自组织形成。

定义 2. 超级节点代理 ASP(Agent Super-Peer)。超级节点代理是联系各个服务组的纽带,逻辑上位于网络最上层。

定义 3. 超级节点 SP(Super-Peer)。超级节点是服务组内推选的在线时间长,高性能的节点。

定义 4. 普通节点 OP(Ordinary-Peer)。普通节点是除了超级节点代理和超级节点之外的节点,是网络的终端用户。

定义 5. 服务能力。节点 I 对节点 J 的服务能力与节点 I 的处理能力(P_i)、信誉度(R_i)、资源数量(N_i)、通信带宽(Cb_i)、平均在线时间(T_i)及节点 I 与节点 J 的内容相似度(S_{ij})正相关,与节点 I 到节点 J 的物理距离(D_{ij})负相关。

设七元组 C_i(P_i, R_i, N_i, CB_i, T_i, S_{ij}, D_{ij})为节点 I 的异构性集合,对应权值为 (a, b, c, d, e, f, j), 则节点 I 对节点 J 的服务能力 SA_{ij} 为:

$$SA_{ij} = (aP_i + bR_i + cN_i + dCB_i + eT_i + fS_{ij} - jD_{ij}) \quad (1)$$

其中权值 (a, b, c, d, e, f, j) ∈ (0,1),大小根据节点自身的需要由节点自己决定,并且有 a + b + c + d + e + f + j = 1。

定义 6. 相似度。节点 I 与节点 J 的内容相似度 S_{ij} 刻画了节点 I 所共享资源与节点 J 所共享资源的相似程度。相似度可以用多种函数来表示,通常采用余弦相似度函数,其计算公式为:

$$S_{ij} = \frac{\sum_k w_{ik} w_{jk}}{\sqrt{\sum_k w_{ik}^2} \cdot \sqrt{\sum_k w_{jk}^2}} \quad (2)$$

其中 W_{ik} 为节点 I 共享资源向量空间的第 k 个向量, W_{jk} 为节点 J 共享资源向量空间的第 k 个向量。s_{ij} ∈ (0,1), S_{ij} 的值越大,表明节点 I 与节点 J 所共享的资源越相似。

2.2 模型的体系结构

节点根据服务能力自组织形成服务组,并从中选出服务能力最强的节点作为超级节点管理该组,再引

