

基于社会学原理的 P2P 网络模型 REC^{*})

肖卫东 唐九阳 汤大权 张维明

(国防科技大学信息系统与管理学院 长沙 410073)

摘要 集中式的 P2P 网络模型规模受限,完全分布的 P2P 网络模型往往混乱低效。借鉴社会学的相应原理,类比人类社会的聚群,针对节点能力的异构性,将系统中的节点进行分层;针对节点参与资源共享的不同动机,对拓扑连接进行分类,按用户需求和共享目的组织网络拓扑,提出一个以社区为基本逻辑管理单位的分层分布式 P2P 网络模型 REC(Resource demand based Community)。分别描述了 REC 的体系结构和相关协议,理论分析和仿真实验表明 REC 模型是合理有效的,为有效资源组织以及高效资源搜索奠定了基础。

关键词 对等网,社会学原理,社区,拓扑连接

A P2P Network Model Based on Socialization Theory

XIAO Wei-Dong TANG Jiu-Yang Tang Da-Quan ZHANG Wei-Ming

(School of Information System and Management, National University of Defense Technology, Changsha 410073)

Abstract Centralized P2P network model is not suitable for the large-scale network, while the pure distributed P2P network model is inefficient. According to the corresponding sociology principle, the layered and distributed P2P network model, named REC, is proposed by referring to the human society community. REC endows high capability peers with more responsibilities to deal with the problem of peer capability heterogeneity, which enhances the performance of the network. REC differs itself from existing P2P network model in that, its topology connections classification mechanism enables the topology construction according to peers' resources and query requirements, reflecting the altruism and selfish nature of peers. The following analysis shows that the REC is reasonable and effective, providing a foundation to effective resources organization and efficient search.

Keywords P2P, Sociology principle, Community, Topology connection

1 引言

P2P(Peer-to-Peer)是在互联网上实施网络计算的一种新的计算模型。P2P 打破了传统的 Client/Server(C/S)模式,网络中的每个节点是逻辑对等的,拥有对等的功能与责任,每个节点既充当服务器,为其它节点提供资源,同时也享用其它节点提供的资源。随着越来越多节点的加入,P2P 网络提供的资源和功能日益丰富。为了满足系统的可扩展性,即资源规模不断扩大、应用不断增长要求以不降低网络的整体性能为目的,其中网络模型是决定系统性能的关键。

目前,非结构化 P2P 网络的模型主要分为 3 类。以 Napster^[1]为代表的集中式 P2P 网络模型利用中央服务器负责目录管理的服务会因为受服务器的限制,存在服务质量无法提高和单点失效的问题,不适合网络规模化发展;完全分布的 P2P 网络模型,如 Gnutella^[2],由于没有中央服务器,无序的网络结构缺乏对资源的有效组织,在资源搜索时存在消息泛滥的问题,导致系统的资源利用效率低下以及系统的可扩展性较差;混合 P2P 网络模型,如 KaZaA^[3]采用基于超级节点的两层结构,虽然考虑了节点能力的异构,但没有从节点参与共享网络动机的角度区分节点间的拓扑连接,未能有效发挥分层结构提升系统性能的潜能。

针对上述问题,借鉴社会心理学和社会行为学的相应原理,类比人类社会的聚群行为,提出一种按节点的资源需求和共享目的组织的 P2P 网络模型(REC)。介绍了 REC 的体系结构,并详细描述了相关的协议,最后通过理论和仿真实验对 REC 进行了分析。

2 社会学依据

社会心理学研究表明,趋群性是人类的本性之一,人类行为的一个重要特点就是趋向于合群^[4]。社会生产方式越发达,生活方式越先进,各类人员的趋群意识和结群倾向就越强烈。聚群成为人类社会生活中最常见、最普遍的社会现象,其影响已深入到社会政治生活、经济生活、文化生活和家庭生活等各个社会生活领域。结合社会行为学分析,这种自发的、无拘无束的聚群现象是人的有目的、有意识的活动,是根据人的意志来进行的。个人出于减少对周围环境的不确定性的需要而交换信息,从而及时调整自身的行为以适应变化了的环境。个人这种不受通常行为规范所指导的、自发的、无组织的、无结构的群体行为方式称为聚群行为。

一般情况下,聚群需要形成聚群动机,聚群动机驱使聚群行为。人的行为无一不是由动机所引导、维持和导向,而动机又是以人的需要为基础。建立在个人的行为模式上,聚群行

^{*}基金项目:国家自然科学基金资助(60172012),湖南省自然科学基金资助(03JJY3110)。肖卫东 博士,教授,研究方向:信息管理、信息系统集成,对等计算;唐九阳 博士,讲师,研究方向:对等计算、信息集成、知识管理;汤大权 副教授,研究方向:信息资源管理、信息检索;张维明 博士,教授,博士生导师,研究方向:信息系统理论、人工智能、智能决策技术。

为也是在相互一系列心理活动的支配下发生的。由此可见, 聚群心理是聚群行为的内在动力, 聚群行为是聚群心理的外在表现^[5]。比如一个对数学理论有着浓厚兴趣的人, 可能更关注与数学相关的资源; 而对喜欢体育的人, 可能更加关注新闻信息。因此, 按照生物学中“人以群分, 物以类聚”的原理, 有着共同兴趣爱好的人, 相互之间容易建立关系。

将人类社会的聚群心理和聚群行为反映到 P2P 系统中, 则体现为节点具有利己和利人的双重动机, 一方面提供共享资源, 另一方面也有自身的资源需求, 而节点不同的动机将导致节点在网络中采取不同的行为, 相应对系统性能带来很大影响^[6~8]。从资源合理组织的视角考虑, 对资源相似节点进行聚类, 可以方便其它节点的资源查找; 另一方面, 根据节点自身的需求, 尽可能把它纳入到最能满足其需求的资源提供节点聚类中, 可以提高节点自身的搜索效率, 从而有效提升系统的整体性能。

3 REC 体系结构

REC 模型的体系结构如图 1 所示。网络在逻辑上划分成若干社区 (community), 每个社区包含若干节点。由于网络中节点存在很强的异构性, 每个节点提供的处理能力和连接带宽并不是相等的, 而且差距很大, 简单起见, 本文只考虑了带宽异构性的情况, 对所有节点依据各自的带宽性能实现分级, 划分为服务节点 (service peer) 和叶节点 (leaf peer) 两类, 其中每个节点只属于一个社区, 如 Peer_{2a} 属于社区 A, Peer_{3c} 属于社区 C。REC 中, 叶节点与服务节点的拓扑连接分为资源连接 (resource-oriented connection) 和需求连接 (demand-oriented connection) 两类, 服务节点与服务节点之间的拓扑连接通过服务节点连接 (service peer connection) 建立, 服务节点连接是不同社区之间沟通的桥梁。

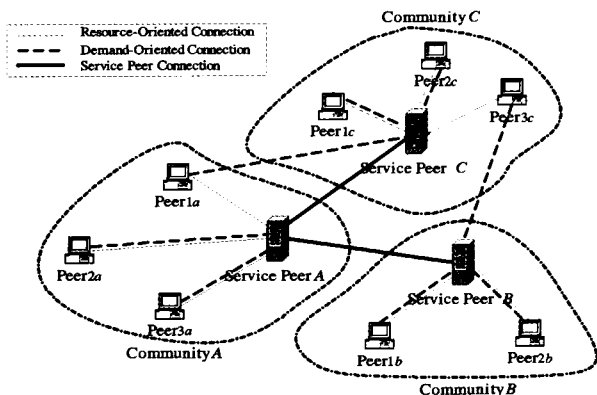


图 1 REC 模型的体系结构

3.1 服务节点

服务节点由带宽能力强 (符合一定带宽标准)、能够长时间稳定工作的节点担任, 逻辑上位于社区的中心, 与社区内叶节点距离最短, 是社区间的接口, 维持一个所辖社区内全部叶节点的局部资源索引, 并代替所属叶节点处理资源请求。每个服务节点都维护以下信息: 社区内叶节点的地址和资源集合; 相邻服务节点的地址集合; 查询统计数据, 如社区内查询消息的数量、查询命中消息的数量、查询命中消息的跳数等。

由于服务节点可能成为社区的瓶颈, 它的瘫痪将造成社区内的部分叶节点不可用。根据文[9]中的设计原则, 可以采取冗余备份的机制, 在每个社区拥有一个服务节点的情况下,

再增加冗余的准服务节点, 在当前社区的服务节点失效时, 由准服务节点取代失效服务节点, 提高社区的稳定性和可靠性。

3.2 叶节点

叶节点类似 C/S 模式中的客户机, 带宽性能较弱, 一个叶节点至少与一个服务节点连接, 叶节点之间不存在逻辑上的连接。叶节点向服务节点报告其共享资源索引, 它们需要保存自身拓扑连接的服务节点的地址集合, 同时维护查询统计数据, 如各社区查询消息的数量、查询命中消息的数量、查询命中消息的跳数等。

3.3 拓扑连接

综合考虑节点的共享资源和资源需求对网络模型的影响, 将叶节点与服务节点的拓扑连接分为资源连接和需求连接两类。

假定系统中的叶节点集合为 $P = \{P_i | 1 \leq i \leq N\}$, 服务节点集合 $SP = \{SP_j | 1 \leq j \leq M\} (M \ll N)$ 。下面给出相应定义。

定义 1 (资源连接) 叶节点基于自身共享资源的资源类型与服务节点的连接称为资源连接, 记资源连接映射为 $L_R: P \rightarrow SP$, L_R 为单射且可逆, 那么服务节点 SP_j 的资源连接节点集合 $L_R^{-1}(SP_j) = \{P_i | L_R(P_i) = SP_j\}$ 。

定义 2 (需求连接) 叶节点基于自身对资源的需求与服务节点的连接称为需求连接, 记需求连接映射为 $L_D: P \rightarrow SP$, L_D 同样为单射可逆, 那么服务节点 SP_j 的需求连接节点集合 $L_D^{-1}(SP_j) = \{P_i | L_D(P_i) = SP_j\}$ 。

3.4 社区

定义 3 (社区) 系统中服务节点 SP_j 以及 SP_j 的所有资源连接节点组成的逻辑覆盖网称作社区。下文中我们将服务节点 SP_j 管理的社区称作社区 SP_j , 记为 $C(SP_j)$, 用图表示为 $C(SP_j) = (N, E)$, 其中节点集合 $N = \{SP_j \cup L_R^{-1}(SP_j)\}$, 边集 $E = \{e | \forall P_i \in L_R^{-1}(SP_j), \text{有 } e \text{ 连接 } P_i \text{ 和 } SP_j\}$ 。

定义 4 (社区连接) 不同社区间服务节点的连接称为社区连接, 记社区连接映射为 $L_C: SP \rightarrow SP$ 。

社区概念与人类社会中的聚群极为契合。类比人类社会的聚群结构, 可将整个 P2P 信息共享空间按兴趣领域划分为若干社区, 社区就是由一些具有共同兴趣的节点所构成的一个虚拟实体, 即构成互联网中的一个逻辑覆盖网, 使信息按新方式又一次集中。通过对信息资源的重组, 以期建立开放的及专业化的社区使兴趣相同的节点相互合作, 简化信息资源的存储、查找和使用, 从而提高资源利用率。

4 协议描述

4.1 节点加入、退出协议

节点的加入 节点初始加入网络时需要了解网络中至少一个节点的地址信息, 并以其为引导节点加入网络。节点根据自身的带宽性能确定节点的身份, 分为叶节点和服务节点两种情况。

如果新加入的是一个叶节点 P_i , 则具体过程如下:

(1) P_i 发送 Connect 消息给引导节点 n , 如果 n 不是服务节点则转(5)。

(2) 设引导节点的服务节点为 $SP_j (n \text{ 或 } L_R(n))$, 节点 SP_j 判断自身拥有的叶节点数, 假设 SP_j 最大可以连接叶节点的容量为 K_p , 倘若当前容量小于等于 $(K_p - 2)$, 那么, P_i 可以直接加入到 SP_j 的叶节点集合中, 转(4); 如果 SP_j 的当前容量已满, 则选择 SP_j 的一个邻居服务节点, 假设为 SP_m , SP_j 将 P_i 的 Connect 消息路由到节点 SP_m 。

(3) 节点 SP_m 重复(2)的步骤,直至 P_i 到达服务节点 SP_i 并成功加入;如果加入消息被转发的次数达到了系统设定的最大值,则将 P_i 设置为引导节点 n 的一个临时节点,在一定时间间隔后重新发起加入操作,这种情况发生的概率很小。

(4) P_i 加入到服务节点(设为 SP_j)管辖的社区中, P_i 和 SP_j 间建立起两条逻辑上的双向连接,默认初始状态下叶节点的资源连接和需求连接为同一服务节点, SP_j 进行相应的信息更新,将叶节点数目增加 2,并增添 P_i 的资源索引,叶节点加入过程完成。

(5) n 将 P_i 的 Connect 消息发送给它的资源连接服务节点 $L_R(n)$,然后转(2)。

服务节点(设为 SP_m)的加入过程与叶节点类似。在定位到引导节点的服务节点 SP_j 后,判断 SP_j 可以连接的服务节点容量(设为 K_{SP})是否已满,如果当前容量小于等于 $(K_{SP}-1)$, SP_m 和 SP_j 之间建立连接,各自将服务节点数目加 1,同时增添彼此的地址信息。

REC 中,节点根据自己了解的系统中其它节点的知识自主地建立并维护与其它节点的邻接关系。为了适应节点资源的动态变化,叶节点定时向其服务节点发送更新消息,服务节点则定时向邻居服务节点发送更新消息。

节点的退出 同样分别考虑叶节点和服务节点两种情况。

叶节点在正常退出网络时,应分别向其资源连接和需求连接的服务节点发送注销报文,在收到注销报文后,资源连接的服务节点删除该叶节点的资源索引,断开与该叶节点的连接,需求连接的服务节点则直接断开与叶节点的连接即可。此外,要对服务节点的信息进行更新,资源连接和需求连接的服务节点的叶节点数目各减 1。对于叶节点发生故障而退出网络的情况,REC 采用探测的方法。即当叶节点在规定时间内未发出任何更新,其服务节点向该叶节点发送超时探测消息,若不能获得响应,则认为该叶节点故障,此时,将像处理叶节点正常退出那样执行相应的更新。

当服务节点正常退出网络时,它首先定位并通知退出以后将代替它的服务节点,新的服务节点搜集退出服务节点的相应信息,并通知原有服务节点的资源连接和需求连接叶节点以及原有服务节点的邻居服务节点,在新的服务节点取代退出服务节点进入正常工作以后,原有服务节点断开自己和其它节点的连接并退出系统。若服务节点因故障或异常退出网络,其邻居服务节点探测到该服务节点未能定期发出更新,则通知系统激活准服务节点,取代失效的服务节点。

4.2 路由搜索协议

对于给定的资源搜索请求,叶节点根据请求内容与自身维护资源的语义相关性决定搜索请求的路由方向(资源连接或需求连接的服务节点)^[10],服务节点一方面查询维护的局部资源索引,命中则将共享此资源的节点信息返回给请求节点,另一方面采用一定的策略向它的邻居服务节点转发资源搜索请求。由于资源搜索请求只在服务节点之间进行,因此减少了叶节点带宽的消耗,大大降低了通信数据量。一旦叶节点发现待发送搜索请求的服务节点失效,则自动缺省将搜索请求发给另一拓扑连接的服务节点。值得注意的是,为了减少系统负载,叶节点只选择一个服务节点发送搜索请求。

5 REC 的分析

5.1 合理性

以社区为基本逻辑管理单位的 REC 是一种集分布式运算及集中式管理为一体的混合 P2P 模式。和集中式结构 Napster 相比,REC 把中心服务器的负载分散到各个服务节点,使得单个服务节点不会因负载过重而成为潜在的系统瓶颈。和全分布式 Gnutella 系统相比,采用高性能节点形成网络的支柱,引入层级结构,扩大了搜索范围,提高了搜索速度,带来资源利用率的提高。

KaZaA 中基于超级节点的两层结构是与本文提出的 REC 相似的 P2P 网络模型。与 KaZaA 不同的是,REC 中拓扑连接分类的引入,体现了节点服务它人和满足自身的双重视角,为网络中的有效资源组织奠定基础。此外,拓扑连接分类使得节点共享资源的变化将独立于资源需求的变化,增强了系统拓扑的稳定性,适应动态自治的环境,而且当叶节点某一拓扑连接的服务节点失效时,可以通过另一拓扑连接实现信息的交流,系统的鲁棒性增强。当然,该处理是以牺牲一定的存储空间为代价的。对于每个叶节点而言,只需维护自身的资源以及两个拓扑连接的服务节点地址信息,存储开销较小,存储开销主要集中在每个服务节点上,服务节点需要保存社区内所有资源连接叶节点的资源索引、资源连接和需求连接叶节点的地址信息、相邻服务节点的地址信息,相比 KaZaA 的拓扑结构中每个节点只连接一个超级节点的情况,服务节点的存储空间增加了一个拓扑连接的地址信息,但是该地址信息相对资源索引来说,存储空间极小,存储开销是可以接受的。

5.2 协议综合性能

采用仿真实验对协议的综合性能进行分析。实验是在一台 PC 机上完成的。配置为 CPU P4 2.0GHz,内存 512MB,操作系统 Windows Professional 2000。整个 P2P 网络由 15 个服务节点和 1485 个叶节点构成,网络拓扑结构由独立的算法产生,服务节点之间的连接满足幂规律($\tau=3$)。实验中随机提取了波音公司 Web 日志数据^[11]中的 1500 个节点和相应的 43240 个文档,每个文档通过一个唯一的序号加以标识。本文假定只要某个被搜索的节点存储了查询所针对的文档,那么搜索成功。

从搜索成功率和搜索路径长度两方面将 REC 与 KaZaA 进行对比分析。搜索成功率是评价网络性能的重要指标,我们将预定时间内结果返回的数量占系统当时所有在线节点维护答案总数的比率称为节点的资源定位率,那么搜索成功率可以定义为系统在一个预定的时间内,所有请求的资源定位率的平均值,用式(1)计算:

$$\text{搜索成功率} = \frac{\text{资源定位率之和}}{\text{资源请求的总数量}} \quad (1)$$

搜索路径长度是决定网络传输时间的直接因素,是衡量网络性能的另一重要指标。搜索请求到达目标地址所经过的跳转数越少,则搜索请求从发出到得到结果所需的延迟时间就会相对较小,从而导致网络性能较优。因此搜索路径长度可以定义为预定时间内所有请求的查询跳数的平均值,用式(2)计算:

$$\text{搜索路径长度} = \frac{\text{资源请求的查询跳数之和}}{\text{资源请求的总数量}} \quad (2)$$

为了描述的方便,我们将连续仿真时间中从初始观测点开始的每一段固定时间间隔称为轮次(round),每个轮次中各节点只给定一个查询动作。在仿真中,我们将轮次设为 21s。

(下转第 67 页)

$$-d^{-k-1}] = 0$$

经整理, 上式等价于:

$$d^{1-k} - (1-d)^{1-k} + 1 - 2d = 0$$

当 $d=1/2$ 时, 满足上式;

当 $1/2 > d > 0$ 时,

$$d^{1-k} - (1-d)^{1-k} + 1 - 2d > 0;$$

当 $1 > d > 1/2$ 时,

$$d^{1-k} - (1-d)^{1-k} + 1 - 2d < 0.$$

显然, 在 d 的取值范围内, 当且仅当 $d=1/2$ 时, 满足

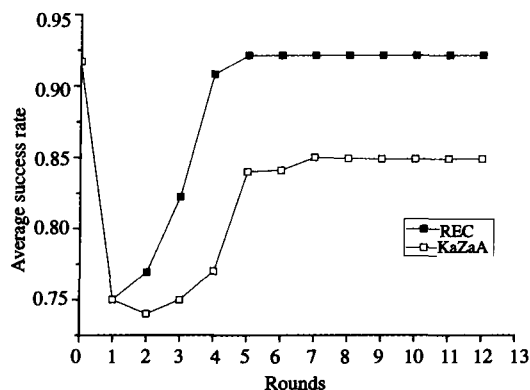
$$d^{1-k} - (1-d)^{1-k} + 1 - 2d = 0$$

结合极值与导数的关系, $d=1/2$ 时, $\frac{2(d(-d^2))^{-k}}{d^{-k} + (1-d)^{-k} - 1}$

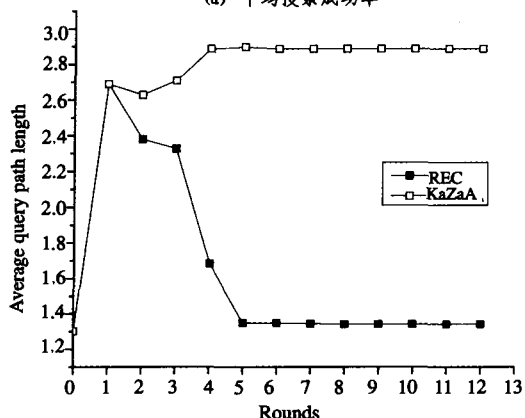
最大。

(上接第 40 页)

采用文[10]的 P2P 网络自组织构造方法 SOMA, 对于稳定后的同一个网络拓扑, 从 1500 个节点中随机选择 500 个叶节点, 该 500 个叶节点发出不同于自身共享资源的搜索请求(模拟资源需求变化), 其余 985 个叶节点还是针对自身资源查询, TTL(Time To Live)值设置为 5。我们在 REC 的基础上采用每个叶节点只连接一个服务节点模拟 KaZaA, 每次实验进行 12 轮次的调整, 数据为 10 次仿真实验的平均, 图 2 展示了 REC 和 KaZaA 随拓扑变化性能的变化。



(a) 平均搜索成功率



(b) 平均搜索路径长度

图 2 REC 与 KaZaA 性能的比较

轮次 0 表示初始稳定网络拓扑下的性能。图中可以发现, 由于部分节点的资源需求发生改变, 在轮次为 1 时, REC 和 KaZaA 的搜索性能同时恶化, 而且保持一致, 这是因为 SOMA 导致的拓扑变化存在一定的滞后, 此时节点的拓扑还未发生改变。在随后的过程中, REC 性能逐渐优化, 这是由于 REC 中节点通过需求连接建立与所需的资源的连接, 同时保持它原来所属的社区并不变化。相对而言, 在缺少需求连接的情况下, 图 2 中 KaZaA 的平均搜索成功率从第 3 个轮次开始提高, 平均搜索路径长度始终呈上升趋势。这是因为 KaZaA 中节点逐步脱离共享资源社区, 加入到查询需求的资

源聚集社区, 虽然提高了自身的搜索性能, 但由于节点的移动影响了原有共享资源的聚集, 破坏了资源社区结构, 从而降低了其它节点针对节点共享资源的搜索性能, 相应导致平均搜索路径长度的增长。

从图 2 我们可以观察到, 基于资源连接和需求连接的 REC 整体性能优于 KaZaA, 这是因为 REC 可以通过网络自组织构造, 一方面将具有相似资源的节点借助资源连接松散地聚合在一起, 增强搜索请求传播的针对性; 另一方面, 借助需求连接建立的捷径, 与资源提供节点保持逻辑拓扑上的相邻, 减少社区间路由转发的次数, 实现节点快速获取所需资源以及被其它节点检索。相比 KaZaA 下或者通过相似资源节点的聚类, 或者将最有可能包含查询结果的节点保持为邻居节点, REC 从利己和利人的双重角度充分发掘社区结构的潜能, 有效提高 P2P 系统的整体性能, 在动态自适应方面更具优势。

结论 REC 结合了传统集中式网络易于管理与分布式网络具有良好的区域自治、负载均衡以及健壮性的优点, 针对节点能力的异构性, 将系统中的节点进行分层, 赋予高性能节点更多职责, 能够利用节点的差异提高网络的性能; 针对节点参与资源共享的不同动机, 对拓扑连接进行分类, 按节点资源需求和共享目的组织网络结构, 体现了节点服务它人和满足自身的双重视角, 为网络中的有效资源组织、高效资源搜索奠定基础。理论分析和仿真实验表明 REC 模型是合理有效的。

参考文献

- 1 Napster. 2003. <http://www.napster.com/>
- 2 Gnutella. 2003. <http://www.gnutella.com/>
- 3 LTD S N. KaZaA media desktop, 2002. <http://www.kazaa.com>
- 4 时蓉华. 社会心理学. 浙江: 浙江教育出版社, 1998
- 5 威尔逊. 论人的天性. 贵阳: 贵州人民出版社, 1987
- 6 Shneidman J, Parkes D. Rationality and self-interest in peer to peer networks. In: Proceedings of the Second International Workshop on Peer-to-Peer Systems (IPTPS). Berkeley: Springer Verlag Press, LNCS 2735, 2003. 47~52
- 7 Vassileva J. Motivating participation in peer to peer communities. In: Proceedings of the Workshop on Emergent Societies in the Agent World (ESAW'02), Madrid, Springer. Verlag LNAI 2577, 2002. 141~155
- 8 Cheng R, Vassileva J. User motivation and persuasion strategy for Peer-to-peer communities. In: Proceedings of 38th Hawaii International Conference on System Sciences (HICSS' 2005). Hawaii, USA, 2005. 193~202
- 9 Yang B, Garcia-Molina H. Designing a super-peer network. In: Proceedings of the 19th International Conference on Data Engineering (ICDE'2003). Bangalore, India, 2003
- 10 唐九阳, 张维明, 肖卫东, 汤大权. 类人类社会基于社区的对等网自组织构造. 计算机研究与发展, 2006, 43(8): 1383~1390
- 11 Davison B. Web caching and content delivery resources. <http://www.web-caching.com/traces-logs.html>