

一种基于覆盖网络拓扑的无结构P2P主动复制策略*

冯国富^{1,2+}, 张金城¹, 顾庆², 陆桑璐², 陈道蓄²

¹(南京审计学院 信息科学学院,江苏 南京 210029)

²(南京大学 计算机科学与技术系,江苏 南京 210093)

An Overlay Topology Based Proactive Replication in Unstructured P2P Systems

FENG Guo-Fu^{1,2+}, ZHANG Jin-Cheng¹, GU Qing², LU Sang-Lu², CHEN Dao-Xu²

¹(School of Information Science, Nanjing Audit University, Nanjing 210029, China)

²(Department of Computer Science and Technology, Nanjing University, Nanjing 210093, China)

+ Corresponding author: Phn: +86-25-58318603, E-mail: fgfmail@dislab.nju.edu.cn

Feng GF, Zhang JC, Gu Q, Lu SL, Chen DX. An overlay topology based proactive replication in unstructured P2P systems. *Journal of Software*, 2007,18(9):2226-2234. <http://www.jos.org.cn/1000-9825/18/2226.htm>

Abstract: The overlay network of unstructured P2P system is neither regular network, nor pure random network. The peers are usually not completely equivalent. They usually play different roles in the overlay network. This paper firstly investigates the relation among the degree distribution, the access frequency mode and the success rate, and then presents an optimal degree distribution model in terms of the popularity of data items. Finally, a feasible proactive replication is proposed to reach the expected degree distribution. The simulation shows that the proactive replication can improve the performance of the unstructured P2P.

Key words: unstructured P2P; proactive replication; distribution of degrees; replica efficiency

摘要: 无结构 P2P 覆盖网络并非规则网络,也非纯粹的随机网络,结点在拓扑结构中体现出非对等性,在接收查询消息的数量上具有非均衡性.研究了结点连接度分布、数据流行程度与搜索成功率之间的关系,并针对数据的不同流行程度给出了结点连接度的最优分配模型.最后给出了一种实现最优结点度分配的主动复制策略.实验结果表明,基于拓扑信息的主动复制是一种提高无结构 P2P 搜索性能的可行方法.

关键词: 无结构 P2P;主动复制;结点度分布;副本效率

中图法分类号: TP393 文献标识码: A

P2P 作为一种全新的分布式计算模式,受到了学术界和 Internet 用户群的广泛关注.P2P 系统中的结点既作为客户端发出请求,又作为服务器端提供服务,消除了传统 C/S 模式中的单点失败和性能瓶颈问题,具有更好的

* Supported by the National Natural Science Foundation of China under Grant Nos.60573106, 605731321 (国家自然科学基金); the National High-Tech Research and Development Plan of China under Grant No.2004AA112090 (国家高技术研究发展计划(863)); the Colleges Oriented Provincial Natural Science Research Plan of Jiangsu Province of China under Grant Nos.07KJD520052, 06KJD520090 (江苏省高校自然科学研究计划)

Received 2006-02-21; Accepted 2006-08-21

容错性和可靠性.

目前的 P2P 系统主要分为两类:结构化 P2P 和无结构 P2P.结构化 P2P 在结点之间添加确定的连接,构造规则的覆盖网络,拥有严格的资源存放规则.结构化 P2P 效率高,但却破坏了 P2P 系统结点自治性,不能适应网络结点的高度动态性.根据元数据管理方式,无结构 P2P 体系结构通常分为集中式、全分布方式和混合式.集中式和混合式通过中心服务器或超级结点提供信息服务,因此容易导致单点失败和性能瓶颈.针对全分布体系结构的搜索算法都是通过查询消息在结点之间的转发、匹配来实现的,搜索算法的性能主要取决于查询消息路由策略和数据复制策略.

依靠查询消息遍历结点的搜索方式较为简单,但其搜索效率低,网络开销大.目前,提高无结构全分布体系结构 P2P 系统性能的方法主要有两种:一种是通过收集、交换结点信息来提高路由效率;另一种是通过大量数据复制来提高命中效率.

数据复制是提高 P2P 搜索性能最常用的策略之一.一方面,P2P 系统中的结点具有高度动态性,随时加入和离开系统,数据复制能够提高数据的可获取性;另一方面,数据复制在降低用户访问延迟、消除热点瓶颈、提高系统可靠性等方面都是显著的.

在当前全分布的无结构 P2P 中,复制策略主要有 3 种:宿主复制、随机复制和路径复制.它们都是为了提高搜索命中率,降低响应时间,在搜索过程中动态地、盲目地、后发性地进行数据复制,只考虑副本数量上的差异,并不考虑存放位置的差异.本文提出了一种基于自组织网络拓扑进行主动数据位置交换以提高 P2P 系统性能的策略.主动复制考虑了覆盖网络结构对搜索性能的影响,能够适应热点变化,提高存储空间使用效率,从而提高搜索性能.

本文第 1 节回顾相关工作.第 2 节分析拓扑结构对搜索的影响,给出主动复制策略.第 3 节给出模拟实验和性能分析.第 4 节给出结论.

1 相关工作

依靠查询消息遍历结点的搜索方式简单、有效,适应 P2P 系统的高度动态性,得到了广泛的研究和应用.Gnutella^[1]是最早采用查询消息广播实现网络资源发现的系统.Gnutella 能够较好地适应结点的动态加入和退出,但会产生大量冗余消息,占用网络带宽,不具有可扩展性.为此,提出了许多算法来提高消息数量的可控性和查询消息的命中率.

1.1 消息转发模式

Flooding 是 Gnutella 采用的消息转发模式,查询源将查询消息转发给所有与之相邻的结点.结点收到查询后匹配本地存储是否满足查询,如果满足,则返回结果;否则继续将查询转发给所有邻居.每个查询消息有一个 TTL(time to live)值,用于控制查询消息转发次数.Modified-BFS^[2]是对 Flooding 的“横向”改进,并非每次都转发给所有邻居结点,而是随机选择 k 个以削弱查询消息扩散.Expanding Ring^[3]是对 Flooding 的“纵向”改进.查询源首先以较小的 TTL 值 Flooding 搜索,如果搜索失败,则增大 TTL 值继续搜索,直到搜索成功或者 TTL 达到某一上限值为止.在 Random Walks^[3]中,查询者生成 k 个查询消息,每个消息在结点之间依次转发传递以匹配查询.这样的每一个消息被称为 walker,如图 1 所示.同样,每个 walker 具有一个 TTL 值来限制最多传递的次数.Random Walks 路由算法具有更大的灵活性,被认为是最好的消息转发模式之一.

对消息转发模式的改进仅仅使得查询消息在数量规模上变得可控,并不能提高搜索效率,因为盲目搜索算法中的每个结点满足查询的概率是相同的,搜索成功的能力取决于数据副本的“密度”以及所遍历结点的数量.为此,有许多工作来提

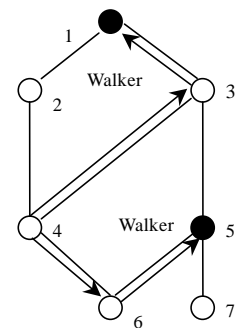


Fig.1 A simple model of overlay network

图 1 一个覆盖网络的简单模型

高查询消息转发的效率,即区分选择最有可能满足查询的结点进行转发,如文献[4]根据 Power-law 拓扑,总是选择具有最大度的结点进行转发.文献[5]采用局部索引算法,通过信息交换提高命中效率.与以往从路由算法和消息模式角度提高搜索性能的方法不同,本文根据覆盖网络特征,区别发现结点的不同“地位”,根据流行程度对数据进行位置交换,主动改变覆盖网络中的数据位置配置结构,提高每次消息转发的效率,从而提高系统性能.

1.2 数据复制

数据复制是提高无结构P2P搜索性能最常用的方法之一.目前的复制策略主要有3种^[3]:宿主复制(owner replication)、随机复制(random replication)和路径复制(path replication).在宿主复制方法中,结点保持成功查询所获取数据的副本(数据本身或者包含数据位置信息的元数据,下同),然后该结点便可以作为服务提供者与其他结点提供数据服务.因此,系统中数据副本的数量与该数据被访问的频率基本成正比关系.在随机复制方法中,数据副本按照一定的概率随机选取结点部署副本.随机复制需要了解覆盖网络结构,而P2P系统中结点仅仅知道其邻居结点信息,因此,随机复制在大部分情况下不容易实现.路径复制是指在搜索成功之后沿消息转发路径逆向沿途复制.Freenet^[6]是采用路径复制的典型系统.Freenet通过大量路径复制和LRU文件副本淘汰使具有相近id值的文件所在结点聚集,形成基于id的路由方向感.但是,路径复制中的大量文件移动会消耗大量的网络带宽.同时,覆盖网络中的一些查询消息“常经”结点通常具有更大的消息压力和热点数据压力.

当前,所有的复制策略都使用LRU或者LFU算法管理存储空间,采用边访问、边淘汰、边复制的模式增加热点数据副本在系统中的分布“密度”,以提高大部分针对热点数据查询的效率,从整体上提高搜索性能.相应地,热点数据的大量副本会“占用”不经常被访问数据的存储空间,甚至将这些数据淘汰出系统,使得针对“冰点”数据的搜索性能下降.本文的复制不是被动地、后发性地进行的,而是自适应地、先发性地发现热点并进行复制,达到数据副本的合理部署.

与仅仅通过调整数据副本数量分配来优化数据副本部署的传统方法^[8]不同,本文通过调整数据副本在网络中的位置来优化数据副本部署.

2 无结构 P2P 中基于拓扑的主动复制策略

P2P 系统首先将结点组织成为应用层覆盖网络,每个结点仅仅掌握邻居结点信息,查询消息在结点之间转发实现查询消息匹配检索.因此,一个结点对整个系统的“贡献”在于所存储副本能够命中查询的能力,体现在该结点能够收到查询的能力以及所存储内容被访问的频率.

2.1 模型假设

假设系统具有 N 个结点和 m 个相异文件对象,每个结点存储相异文件的数量为 c ,系统所有相异文件集合为 O ,所有文件副本的集合为 R .用 q_i 代表系统文件 i 的访问频率,也就是在最近一段时间内对文件 i 的访问次数占总访问次数的比例,则

$$\sum_{i \in O} q_i = 1.$$

假设数据文件最初以随机复制方式、Uniform分配模式进行部署,即每个数据对象拥有相同的数据副本数量,副本具有随机的网络位置.文件 i 所有副本集合用 ρ_i 表示.文件 i 所有副本所在结点连接度之和用 r_i 表示.

假设覆盖网络中结点之间的邻居关系构成一个连通无向图 $G=(V,E)$,其中, V 为结点集合, E 为无向边的集合.当结点加入到系统中时分配以固定的连接数,整个网络维持固定的平均度 \bar{d} ,总结点度数为 D .结点以相同的概率发起查询.采用Random Walks路由算法,walker的个数设定为 w .

2.2 网络拓扑与结点收到查询消息能力

如果将一个文件副本部署在 u 结点,那么该副本的“价值”等于所有结点发起针对该文件的查询,而且查询消息能够遍历到结点 u 的次数.

在基于TTL的搜索算法中,一个结点被遍历的概率与消息模式和网络拓扑相关.在Flooding模式中,只有 u

结点有效覆盖区域以内的结点,即与 u 结点距离小于 TTL 的结点,发起的查询才能到达 u 结点;而且,从 u 结点有效覆盖区域内结点发起的查询也必将能够到达 u 结点.所以,在 Flooding 搜索算法中增加一个副本的“收益”取决于所存放该副本结点的有效覆盖区域内的结点数量.在 Random Walks 模式中, u 结点有效覆盖区域内结点发起的查询将不一定遍历到结点 u ,结点 u 被遍历的概率往往与网络结构、该结点所处的位置,以及查询在网络上的分布相关.如图 1 所示,当结点 4 收到查询并转发,使用 Random Walks 模式,其邻居结点收到该查询消息的概率分别为 $1/3$,即结点 4 度数的倒数.每个结点发起一次查询,一跳之内到达结点 1 和结点 4 的平均消息个数分别为 $5/6$ 和 $4/3$ (虽然查询消息一般不按原路返回,但在结点度数较大时返回次数可忽略).

平均情况下的一次查询,结点 u 能够收到查询消息的概率 M_u ,可以根据 Random Walks 消息路由算法得出.由于结点产生消息的概率相同,按 Random Walks 消息路由算法,消息经过每个连接的概率也近似相等,消息共转发 $w \times TTL$ 次,因此,消息到达节点 u 的概率为

$$M_u = d_u \times (w \times TTL) / D \quad (1)$$

式(1)表明,平均情况下,一个结点收到查询消息的能力与其连接度成正比关系.而一个结点在一次查询中只有一次遍历是有效的,在结点度分布差异不是非常巨大的情况下,本文假设结点收到的消息都为有效消息,忽略在一次查询中多次收到的消息.

2.3 结点连接度和数据访问频率与搜索成功率

一个结点所共享存储空间对系统的“贡献”不仅与共享空间的大小、结点的连接度有关,还与所存储数据的热门程度相关.

首先考虑每个结点只共享 1 个存储空间的情况.

按照发起查询分布规律共发起 S 次查询,则共产生 $TTL \times w \times S$ 个消息,平均每个结点发起查询的次数为 S/N .当为文件 i 副本所在的所有结点仅仅分配 1 个连接度,且令 $T = w \times TTL$ (即每次查询产生的消息总数)时,根据式(1),该结点在 S 查询中收到的有效查询消息数目为

$$EM = T \times S / D \quad (2)$$

也就是说,当为文件 i 副本所在的所有结点仅仅分配 1 个连接度时,所有副本能够命中 EM 次查询,能够平均保证 n 个结点的所有查询,其中

$$n = EM / (S/N) = T \times N / D \quad (3)$$

但是,当为文件 i 的副本所在结点分配第 2 个连接度时,所保证查询的新增结点个数会小于式(3)给出的结果.这是因为部署第 2 个连接度所收到的有效查询消息仅仅对没有保证的结点有效,对于已经被前一个连接度所保证的结点来说是无效的.以此类推,增加的连接度越多,一个新增度所能保证查询的新增结点个数就会越少.用 n_i 代表第 i 个连接度部署之后所能保证查询的结点总个数,逐渐增加文件副本所在结点的连接度,并令 $\Delta = T \times N / D$,则

$$\begin{cases} n_1 = \Delta \\ n_2 = n_1 + \Delta(1 - n_1 / N) \\ n_3 = n_2 + \Delta(1 - n_2 / N) \\ \dots \\ n_i = n_{i-1} + \Delta(1 - n_{i-1} / N) \end{cases} \quad (4)$$

由式(4)可得 d 个连接度部署之后能够覆盖的结点总个数为

$$n_d = \Delta \sum_{k=0}^{d-1} \left(\frac{N - \Delta}{N} \right)^k = N [1 - (1 - T/D)^d] \quad (5)$$

因为结点发起查询的概率相同,所以,部署 d 个连接度后的搜索成功率 SR' 为被保证查询的结点与结点总个数的比值.

$$SR' = n_d / N = 1 - (1 - T/D)^d \quad (6)$$

以上结论是基于系统只有一个数据对象的假设,所有访问都是针对该数据对象;只要收到查询消息,该副本

就一定能够命中.当系统中存在多个数据对象时,一个副本与查询匹配的概率为对该对象的访问频率.所以,数据对象 i 的所有副本对整个系统搜索成功率的贡献为

$$SR_i = q_i(1-(1-T/D)^d) \quad (7)$$

式(7)表明,一个数据对象对整个系统的搜索成功率是其副本所部署结点连接度 d 的函数.结合结点总度数限制(每个结点共享 c 个空间,存储 c 个文件副本,总结点度复用为 $D \times c$),

$$\sum_{i=1}^m r_i = cD \quad (8)$$

由拉格朗日乘法可得结点连接度对数据访问频率的最优分布:

$$\begin{cases} r_1 = cD \left/ \left[(\log_{1-T/D}^{q_1}) \sum_{k=1}^m \frac{1}{\log_{1-T/D}^{q_k}} \right] \right. \\ r_i = \frac{\log_{1-T/D}^{q_i}}{\log_{1-T/D}^{q_1}} r_1 \end{cases} \quad (9)$$

2.4 主动复制策略

式(9)表明,结点度对访问频率的合理分配可以优化搜索性能.式(9)为副本放置提供了依据,在最优情况下,访问频率较高的数据对象分配到较多的结点连接度,相应地,应存储于较多的结点或者连接度较大的结点.但对于全分布的 P2P 系统来说,没有全局信息,结点动态性强,难以精确方式达到数据预期分配.相关工作中一般采用“边访问,边复制”的方法,使热点数据拥有较多副本,相应地可以分配到更多的连接度.本文保持数据副本的数量不变,基于式(9),根据数据的不同热门程度和副本所在结点的不同连接度,对数据所在的位置进行动态交换,以达到结点连接度对数据对象的优化分配.

在查询消息路由算法确定的情况下,结点 u 收到针对文件 i 的查询消息的能力与两个方面的因素相关:一方面如前面所述,与结点在网络拓扑中的位置相关;另一方面,还与文件 i 的副本在周围结点的分布密度相关.这是因为,针对文件 i 的查询一旦在周围结点命中,消息转发便会停止.文件 i 在 u 结点周围的分布密度越高,结点 u 收到关于文件 i 的查询概率就越小.

由于分布式 P2P 系统中没有全局信息,单个结点无法获知每个数据对象的访问频率.这里定义一个副本的效率 E 为一段时间内该副本命中查询的次数与所在结点收到查询总次数的比值.鉴于前面结点发起查询频率相同这一假设,副本的效率低说明文件的访问频率不高,或者周围的副本数目较多;反之亦然.由于副本效率反映了数据的热门程度或者数据副本的稀疏程度,所以副本效率具有相对稳定性.设文件 i 和文件 j 有两个副本分别存储于结点 u 和结点 v ,并分别用 E_i 和 E_j 表示其效率,用 H_i 和 H_j 表示其最近一段时间命中的查询消息次数.结点 u 和结点 v 收到的查询消息数量分别为 Q_u 和 Q_v ,则

$$E_i = H_i / Q_u, E_j = H_j / Q_v \quad (10)$$

由第 2.2 节分析可知,一个结点单位时间内收到的查询数量与其结点连接度成正比,所以

$$H_i = T d_u E_i / D, H_j = T d_v E_j / D \quad (11)$$

每个结点需要维护一个邻居信息列表和资源目录列表.邻居信息列表表项主要包括邻居的 IP 位置信息和连接度信息.邻居之间通过“心跳包”交换信息维护连接.资源列表表项主要包括文件的题目、文件类型、关键字、存储路径等信息,还包括该副本的命中历史记录,即用以计算副本效率的命中查询的次数和记录的起始时间.每当结点的邻居连接发生改变,结点自身和资源数据项的历史信息记录清零.

收到查询消息的邻居结点依据下面的原则确定是否进行数据移动、位置交换.

(1) $E_i > E_j, d_u > d_v, d_u > d_v$ 表明 u 结点在拓扑中处于更重要的位置, $E_i > E_j$ 说明文件 i 比文件 j 具有更高的访问频率,并假设数据访问频率、副本效率具有相对稳定性.如果将两个副本位置互换,则可以命中的次数 $H'_i + H'_j = T d_v \times E_i / D + T d_u \times E_j / D < T d_u \times E_i / D + T d_v \times E_j / D = H_i + H_j$, 所以,不进行数据位置交换将具有更高的提高搜索成功率.

(2) $E_i < E_j, d_u > d_v, d_u > d_v$ 表明 u 结点在拓扑中处于更重要的位置, $E_i < E_j$ 说明文件 j 比文件 i 具有更高的访问频

率.如果将两个副本位置互换,由于访问频率具有相对稳定性,则可以命中的次数 $H'_i + H'_j = Td_u \times E_i/D + Td_u \times E_j/D$, 大于 $H_i + H_j = Td_u \times E_i/D + Td_v \times E_j/D$.所以,如果将文件 j 和文件 i 进行位置互换,则可以得到更高的搜索成功率.当位置互换后,修改副本命中次数记录,以免记录不能反映数据效率,引起不合理的交换,令 $H_i = H_i \times d_v/d_u, H_j = H_j \times d_v/d_u$.

(3) $E_i > E_j, d_u < d_v$.同(2),需要进行数据位置互换,并修改副本命中记录.

(4) $E_i < E_j, d_u < d_v$.同(1),不需要数据位置互换.

如果拓扑修改后的时间太短,副本效率不足以反映数据的访问频率,则忽略.这样,经过持续的数据副本位置互换,最终收敛状态的数据配置将使搜索成功率得到优化.

3 实验结果与分析讨论

3.1 实验环境

实验分别采用了基于BA模型^[9]的Power law拓扑和基于Waxman模型^[10]的随机图作为模拟拓扑. Power-law是一种常见的自演化网络拓扑,Gnutella覆盖网络也表现出了Power-law规律^[11].Waxman是常用来模拟Internet拓扑的模型之一.每种拓扑包括 10 000 个结点,结点平均度为 4.1,两种拓扑中结点个数随结点度数的分布曲线如图 2(a)、图 2(b)所示.设置 200 个相异数据对象.每个结点共享 5 个文件副本的存储空间,同一结点不存在同一文件对象的两个副本.

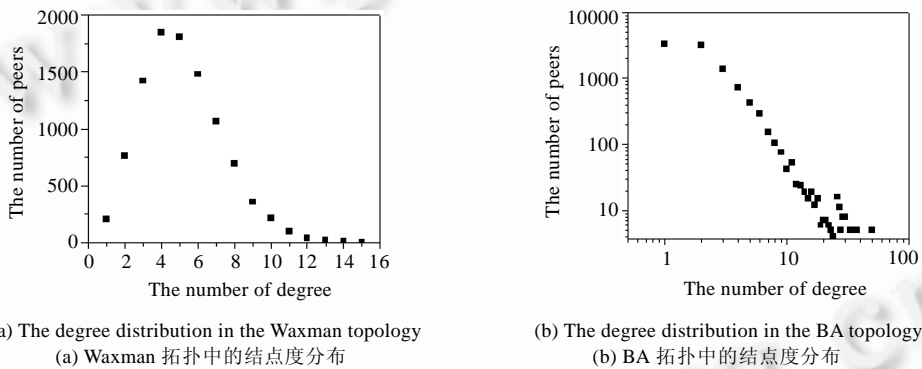


Fig.2 The degree distribution of the topologies in simulations

图 2 实验采用拓扑中的结点度分布

结点以相同的概率发起查询.在 Random Walks 搜索算法中,设置 walker 个数为 3,小于结点平均度数,TTL 设置为 7.以搜索成功率为主要测度来衡量系统性能.相关的实验参数及其缺省值见表 1.

Table 1 The parameters and their default values in simulation

表 1 实验环境参数及其缺省值

Parameters	Value	Parameters	Value
The number of peers	10 000	The number of connections	20 000
Datastore size c	5	The number of data objects	200
Routing algorithm	Random walks	The number of walkers w	3
Power-Law topology	BAmodel	Random graph	Waxman model
Query rate	Zipf ($\alpha=0.92$)	TTL	7

3.2 网络拓扑与结点收到查询消息能力

实验首先验证了在 Random Walks 查询消息路由算法下结点收到查询消息能力和覆盖网络拓扑间的关系.每个 walker 达到 TTL 终止,统计 10 000 次查询中每个结点收到的查询消息数量.理论结果由式(1)给出.从图 3 可以看出,实验结果与理论结果相吻合,具有一致性.

图 4 验证了 Random Walks 路由、副本随机部署策略下不同拓扑中结点度与搜索成功率之间的关系.系统只有一个数据对象,每个步骤开始随机选择 20 个结点添加副本.计算每个步骤的平均搜索成功率、统计副本所在结点的所有连接度,并输出.理论结果由式(6)给出.图 4 显示,实验结果与理论结果在两种拓扑下都相吻合,说明关于网络拓扑与搜索成功率之间确切关系的推理是正确的,结果是可信的.

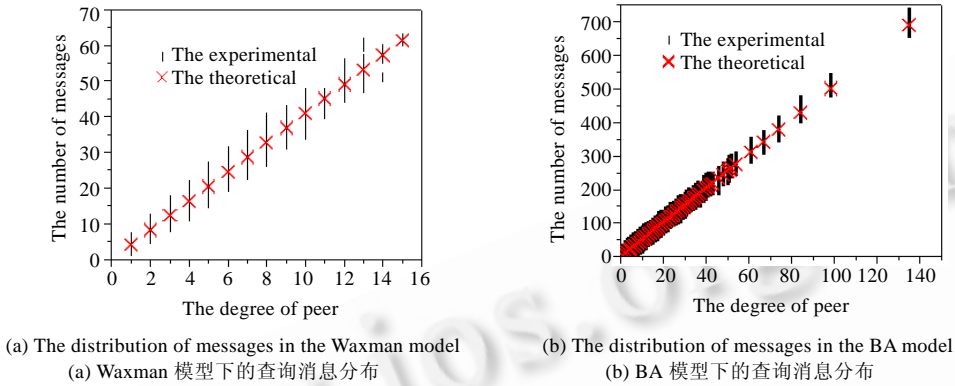


Fig.3 The relation between the number of messages that a peer receives and its degree
图 3 一个结点收到查询消息数量与其结点度之间的关系

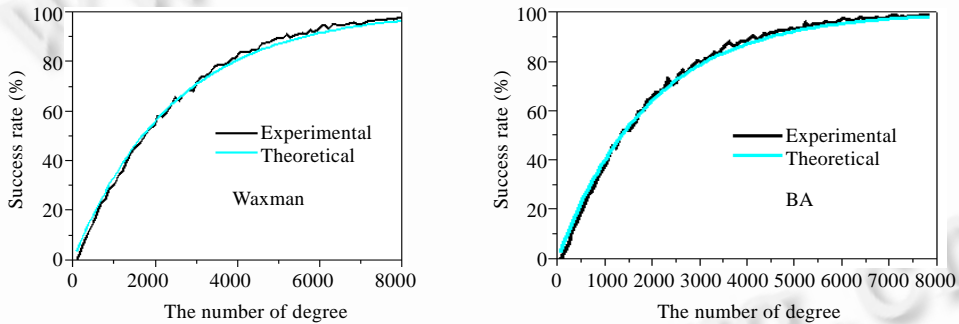


Fig.4 The success rate tendency with the increase of degree

图 4 搜索成功率随分配结点度的变化趋势

3.3 基于拓扑的主动复制策略

接下来考察 Power-law 拓扑下基于拓扑进行主动复制策略的有效性.

图 5 表示出没有复制、宿主复制策略和主动复制搜索性能上的比较.以每秒为时间片,每时间片内随机抽取结点发起 400 次查询,计算每个时间片的平均搜索成功率.在主动复制(proactive replication)的前 150 秒不进行数据交换,即没有数据复制.如图 5 所示,在没有数据复制时,搜索成功率一直维持在 41%左右.宿主复制通过热点数据对冰点数据的替换,使得搜索成功率也逐步提高.但是随着替换的进一步进行,冰点数据逐渐减少以至消失,使得搜索成功率反而下降.宿主复制存在的问题就在于难以维持最佳数据配置状态.当采用主动复制时,搜索成功率提高到 65%左右.这是因为通过数据的位置交换可以使结点度对数据访问频率的分布更接近最优,如图 6 所示.

图 6 是各种策略收敛状态下的结点度对数据访问频率的分布曲线(数据按流行程度大小排序).随机部署策略下每个数据对象分配到的结点平均度基本一样.访问频率较高的数据分配到的结点度与式(9)给出的理论最优值差距巨大,针对这少量结点的大部分查询也不会具有更高的命中机会,整体搜索成功率不能提高.宿主复制收敛过程将冰点数据副本逐渐替换出系统,使得分配到的结点度逐渐减少.相当一部分的数据在系统没有副本

存在,使得命中失败.基于拓扑的主动复制收敛后的度分布与理论最优最为接近.虽然热点数据的结点度数量超过搜索成功率最佳,但可以缩短热点数据的搜索路径长度.大部分的普通数据的结点连接度数量略小于理论最优.

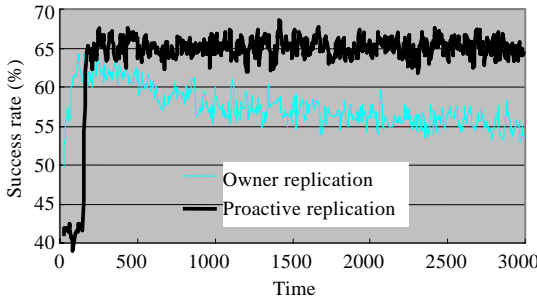


Fig.5 The evolutions of success rate over time with different replication strategies

图 5 不同复制策略下的搜索成功率随时间的演化曲线

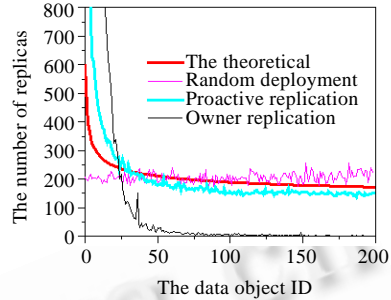


Fig.6 The distribution of degree with different replication strategies

图 6 不同复制策略下结点度对数据的分布

3.4 基于拓扑的主动复制策略动态性能

图 7 是基于拓扑主动复制策略在状态收敛过程中数据位置交换次数的统计结果.采用图 5 中所描述的收敛过程.在收敛初期,需要数据移动再配置的次数较多.2 000 秒,即 800 000 次查询后,每秒交换次数稳定在 13 次左右,相当于每 30 次查询交换一次,每 630 次查询转发发生一次数据移动.因此,基于拓扑的主动复制策略给系统带来的数据移动开销在可接受的范围之内,尤其是仅仅对描述数据位置信息的元数据进行交换时,带来的网络开销会更小.

图 8 考察了主动复制对结点离线的适应性.每秒发起 400 次查询,每秒随机抽取一个结点离线,直至 3 000 个结点离线.计算每秒内的平均搜索成功率.图 8 是主动复制和宿主复制策略搜索成功率随结点离线的分布情况.可以看出,基于拓扑的主动复制策略具有很好的容错性,对结点加入和离开具有很好的适应性.主要原因在于,根据式(1)和式(9),主动复制主要与系统的平均结点度、结点共享空间大小相关,与结点数量无关.而系统的平均度并不随结点离开发生变化,通过基于结点度的数据移动,在结点离线之后达到数据的再配置.所以,系统在结点逐次离线后的搜索成功率基本维持不变.对于宿主复制,结点离线,存储空间减少,系统仅存少量热点数据,对非热点数据访问的命中率再次降低,所以搜索成功率也逐渐降低.

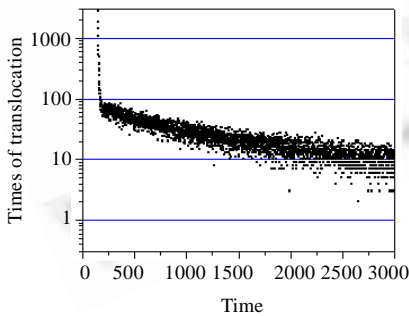


Fig.7 The evolution of translocation times over time in the proactive replication

图 7 主动复制中数据位置交换次数随时间的演化曲线

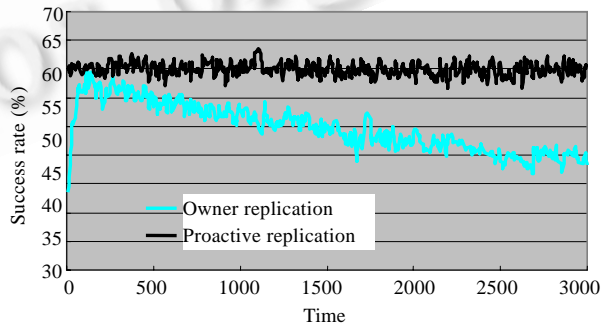


Fig.8 The evolutions of success rate over the gradual leaving of peers

图 8 当结点逐步离线时的搜索成功率演化曲线

4 结 论

无结构 P2P 覆盖网络既非规则网络,也非纯粹的随机网络,结点在拓扑结构中体现出非对等性,在消息转发过程中体现出非均衡性.本文从覆盖网络拓扑的角度研究数据分配和配置,通过数据位置交换、网络结点度对数据的分配,达到突出热点数据的效果,与以往单纯依靠改变数据副本数量的复制策略有所区别.本文首先研究了网络拓扑对结点命中查询能力的影响,然后研究了结点连接度对不同热点程度数据的优化分配问题,给出了结点连接度、数据流行程度与搜索成功率之间的关系.最后给出了相关的主动复制策略.实验表明,基于拓扑信息的主动复制具有自适应性,是一种提高无结构 P2P 搜索性能的可行方法.

References:

- [1] Gnutelliums LLC. Gnutella protocol specification version 0.4. 2006. http://www9.limewire.com/developer/gnutella_protocol_0.4.pdf
- [2] Kalogeraki V, Gunopulos D, Zeinalipour-Yazti D. A local search mechanism for peer-to-peer networks. In: Proc. of the 11th Int'l Conf. on Information and Knowledge Management. New York: ACM Press, 2002. 300-307.
- [3] Lü Q, Cao P, Cohen E, Li K, Shenker S. Search and replication in unstructured peer-to-peer networks. In: Proc. of the Int'l Conf. on Measurements and Modeling of Computer Systems. New York: ACM Press, 2002. 84-95.
- [4] Adamic L, Huberman B, Lukose R, Puniyani A. Search in power law networks. Physical Reviews, 2001,E64:46135-46143.
- [5] Daswani S, Fisk A. Gnutella UDP extension for scalable searches (GUESS) v0.1. 2006. https://www.limewire.org/fisheye/browse/~raw,r=1.2/limecvs/core/guess_01.html
- [6] Clarke I, Sandberg O, Wiley B, Hong TW. Freenet: A distributed anonymous information storage and retrieval system. In: Federrath H, ed. Proc. of the Workshop on Design Issues in Anonymity and Unobservability. Berlin: Springer-Verlag, 2001. 46-66.
- [7] Qin L, Pei C, Edith C, Kai L, Scott S. Search and replication in unstructured peer-to-peer networks. In: Proc. of the 16th ACM Int'l Conf. on Supercomputing (ICS 2002). New York: ACM Press, 2002. 84-95.
- [8] Crucitti P, Latora V, Marchiori M, Rapisarda A. Efficiency of scale-free networks: Error and attack tolerance. Physica A, 2003,320: 622-642.
- [9] Waxman random network topology generator. 2006. <http://www.math.uu.se/research/telecom/software/stgraphs.html>
- [10] Ripeanu M, Foster I. Peer-to-Peer architecture case study: Gnutella network. Technical Report, TR-2001-26, University of Chicago, 2001.



冯国富(1977—),男,山东沂水人,博士,讲师,主要研究领域为分布式计算,协议测试.



陆桑璐(1971—),女,博士,教授,博士生导师,CCF 高级会员,主要研究领域为高性能计算.



张金城(1962—),男,教授,主要研究领域为计算机审计,CIMS.



陈道蔷(1947—),男,教授,博士生导师,CCF 高级会员,主要研究领域为分布式计算.



顾庆(1972—),男,博士,副教授,CCF 高级会员,主要研究领域为分布式计算,软件过程,软件测试.