

结构化 P2P 网络路由算法在网络层的性能评估

于 婧, 胡曦明, 伊 鹏, 汪斌强

(国家数字交换系统工程技术研究中心, 郑州 450002)

摘要: 基于结构化 P2P 路由系统的马尔可夫链模型, 以平均路径长度为指标, 提出了结构化 P2P 网络路由算法在网络层面的性能量化评估方法。该方法解决了结构化 P2P 网络采用基于 overlay 层面的路由算法实现资源查找和定位导致的网络层路由性能无法估量的问题, 能够准确表现结构化 P2P 网络 overlay 层面路由算法的实际网络性能, 为结构化 P2P 网络路由算法的研究提供准确的性能评估机制。

关键词: 结构化 P2P 网络; 路由; 马尔可夫链模型; 平均路径长度

Quantitative Evaluation on Routing Performance on Network-level of Structured P2P Network

YU Jing, HU Ximing, YI Peng, WANG Binqiang

(Engineering and Technology Research Center of National Digital Switching System, Zhengzhou 450002)

【Abstract】 Based on structured P2P routing system model and Markov-chain model, this paper presents a quantitative evaluation on routing performance on the network-level of structured P2P network. Utilizing the average path length as the gist, this method resolves the difficulty induced by the routing algorithms on the overlay-level. It can exactly evaluate the routing performance on the network-level, and provide an evaluation mechanism to the research on the routing algorithms of structured P2P network with exactness.

【Key words】 structured P2P network; routing; Markov-chain model; average path length

1 概述

对等网络(P2P网络)是一种全新的网络模式,网络中的每个节点的地位都是对等的,兼有客户机和服务器的功能。P2P网络打破了传统的(C/S)模式,克服了C/S模式中服务器的瓶颈问题,充分利用了网络中的资源。P2P技术的出现极大地促进了网络资源共享及查询、协同计算以及即时通信的发展。P2P网络按照资源组织与定位方法可以将其简单地分为非结构化P2P网络(unstructured P2P network)和结构化P2P网络(structured P2P network)^[1]。结构化P2P网络因其高效快速的资源定位方式成为P2P网络发展的方向。

结构化P2P网络的主要思想是在IP网络的基础上建立一个overlay网络,通常将这两个层面称之为网络层面(network-level)和overlay层面(overlay-level)^[2],所有有关查找和选路的操作都是基于overlay网络进行的。两层网络之间的对应关系问题,很可能导致overlay层面的两个相邻节点在网络层面相距很远,使得在overlay层面上一个Hop的延迟,反映到网络层上是要更大的延迟。目前,结构化P2P网络的路由算法的研究较多,且都是以overlay层面和网络层面基本一致为前提的,而实际上这两个层面的一致性并没有得到很好的解决。因此,overlay层面的路由在网络层面的性能是很值得研究的,而当前并没有在网络层面上对P2P路由算法的量化评估的方法。

本文提出一种在网络层面上对P2P路由算法进行量化评估的方法。该方法的提出能够有力地验证结构化P2P网络overlay层面的路由算法的实际性能,为结构化P2P网络路由算法的研究提供准确的性能评估机制。本文给出了通用结构化P2P网络模型,该模型抽象了结构化P2P网络众多实现方

式的基本概念,有利于简化描述下一部分的分析。在利用马尔可夫链对结构化P2P网络路由系统建模的基础上,给出了结构化P2P网络路由算法在实际网络层性能的量化评估方法。由于平均路径长度是路由算法的重要参数指标,该方法本质上是对overlay层面的路由算法在网络层面的平均路径长度的计算。利用该方法进行计算模拟,评估在overlay层面的两种路由算法在实际网络也就是在网络层面上的性能,并对结果进行分析说明。

2 通用结构化 P2P 网络路由模型

文献[3, 4]对结构化的P2P网络路由系统定义了一种抽象的模型,用来涵盖CAN, Chord, Pastry以及Tapestry等协议的基本概念。该模型可以使本文的讨论独立于特定的协议。

在模型中,每个网络节点都被分配有唯一的节点标识(nodeId)。资源对象也通过系统定义的Hash函数产生唯一的资源标识(key)。节点标识和资源标识来自同一个id空间。将资源key存储在nodeId与key相等或者相近的节点上,该节点称之为key的根(root)。P2P路由就是将对key的查询消息路由到key的根。为了有效地查找,所有的节点都维护一张节点和节点IP的对应表,也称路由表。另外,每个节点都维护一个邻居集合,该集合由一些在id空间与该节点nodeId相近的节点组成。由于nodeId是随机分配的,因此节点的邻居集合也代表的是所有参与节点的一个随机组合。

基金项目: 国家“863”计划基金资助项目(2005-AA-121210)

作者简介: 于婧(1979-),女,博士研究生,主研方向:对等网络;胡曦明、伊鹏,博士研究生;汪斌强,教授、博士生导师

收稿日期: 2006-07-26 **E-mail:** yujing@mail.ndsc.com.cn

与当前网络选路原则一致，路由是向着更接近目的地的方向进行。结构化 P2P 网络路由系统也不例外，每个节点都是选择最靠近目的地的邻居作为路由的下一跳。当前结构化 P2P 路由系统中，每个途经节点都是以 nodeId 与 key 数值距离的远近(也就是 overlay 层面距离)作为路由选择的标准来选择下一跳的。然而，由于 P2P 网络中网络层面和 overlay 层面的不匹配，因此在 overlay 层面上的距离相近不一定等同于实际网络中的距离相近。这种 overlay 层面上的路由算法在网络层面上不一定会起到积极的作用。这也是本文重点建模、模拟和认证的一个问题。

3 结构化 P2P 网络路由算法量化评估

3.1 P2P 路由系统的马尔可夫链模型描述

文献[5]中给出了一种描述 P2P 系统路由查找的马尔可夫链模型。在稳定的 P2P 网络中，持有请求消息的节点可以看作是一个状态，路由过程可以看作是在不同的状态间的迁移。在 P2P 路由过程中，一个节点根据当前路由表将请求发送至另一个节点。此路由过程可模型化为离散马尔可夫链。

定义随机过程 $\{X_h : h = 0, 1, \dots\}$ ， X_h 为路由过程中到达第 h 跳节点时节点的状态。 X_h 可能为失效状态，或者是消息传送到第 h 跳的节点的 nodeId。消息被转发到目的地或者最终丢弃。这 2 种状态为吸收态，其他状态为非吸收态。

假定 P2P 网络中有 n 个节点，其中，0 为目的地；1, 2, ..., $n-1$ 为源； n 为失效节点。如果消息不能被传送到系统内的任何节点，消息就会被丢弃，转到虚拟的失效节点 n ，则 0 和 n 成为 2 个吸收态。定义 i 为第 i 个节点的状态。则转移概率： $p_{i,j} = P[X_h = i | X_{h-1} = j]$ 。即在第 $h-1$ 跳节点的状态为 j 的情况下，第 h 跳节点的状态为 i 的概率。则转移概率矩阵：

$$P = (p_{i,j})_{(n+1) \times (n+1)} = \begin{pmatrix} 1 & 0 & \dots & 0 & 0 \\ U & Q & & & V \\ 0 & 0 & \dots & 0 & 1 \end{pmatrix}$$

其中， Q 是 $(n-1) \times (n-1)$ 矩阵，是状态 1, 2, ..., $n-1$ 的状态转移矩阵； U 和 V 是 $(n-1) \times 1$ 矩阵。

3.2 路由算法量化评估

平均命中率和平均路径长度是衡量路由算法性能的重要指标。本文采用这 2 个指标给出路由算法量化评估方法。利用马尔可夫链的特性计算平均命中率和从源到目的地的平均路径长度。

(1) overlay 层面路由算法量化评估

定义命中率 $a_{i,j}$ 为一个消息成功地从 i 到吸收态 j 的概率；定义路径长度 $m_{i,j}$ 为对应的跳数。

引理 定义矩阵

$$A = (a_{1,0}, a_{2,0}, \dots, a_{n-1,0})^T \quad M = (m_{1,0}, m_{2,0}, \dots, m_{n-1,0})^T$$

则 $A = (I - Q)^{-1}U$ ， $M = ((I - Q)^{-1}A) \div A$ (对任意 i, j ，定义 $Z = X \div Y$ 为 $z_{i,j} = x_{i,j} \div y_{i,j}$) 平均命中率和平均路径长度： $\bar{a} = \frac{1}{n}(\pi^0 A + 1)$ ， $\bar{m} = \frac{1}{n}\pi^0 M$ 。其中， $\pi^0 = (1, 1, \dots, 1)$ 。引理证明可参考文献[5]。

(2) overlay 层面的路由算法在网络层面的性能量化评估

结构化 P2P 网络的路由计算都是在 overlay 层面上进行的，而实际路由的过程是在网络层面展开的。上述模型只考虑 overlay 层面上的平均路径长度，实际请求经历的路径长度应该是网络层面的路径长度。由于 P2P 网络中网络层面和

overlay 层面之间的不匹配，因此这两个层面的路径长度可能会存在极大的差异。下面考虑网络层面上的平均路径长度，给出 overlay 层面路由算法在网络层面上的以平均路径长度为参量的性能量化评估的表达式。

定理 1

定义 $b_{i,j}$ ($i, j \in (0, 1, \dots, n-1)$) 为网络层面节点 i 和 j 间的实际距离，定义矩阵

$$B = \begin{pmatrix} b_{1,1} & b_{1,2} & \dots & b_{1,n-1} \\ \dots & \dots & \dots & \dots \\ b_{n-1,1} & b_{n-1,2} & \dots & b_{n-1,n-1} \end{pmatrix}, \quad C = \begin{pmatrix} b_{1,0} \\ b_{2,0} \\ \dots \\ b_{n-1,0} \end{pmatrix}$$

为对应的距离矩阵。定义路径长度 $m'_{i,j}$ 为网络层面节点 i 和 j 间的路径长度， $M' = (m'_{1,0}, m'_{2,0}, \dots, m'_{n-1,0})^T$ 为对应的网络层面上的平均路径长度矩阵。则

$$M' = ((I - Q)^{-1}((Q * B)A + U * C)) \div A \quad (\text{对任意 } i, j \text{ 定义 } Z = X \div Y \text{ 为 } z_{i,j} = x_{i,j} \div y_{i,j}; \bar{m}' = \frac{1}{n}\pi^0 M'). \text{ 平均命中率计算同引理。}$$

证明

由引理中矩阵 M 的推导，结合网络层面距离矩阵 B, C 可以得到下面的表达式：

$$a_{i,0} m'_{i,0} = p_{i,0} b_{i,0} + \sum_{\substack{j=1 \\ j \neq i}}^{n-1} p_{i,j} a_{j,0} (m'_{j,0} + b_{i,j})$$

分解组合得到：

$$a_{i,0} m'_{i,0} - \sum_{\substack{j=1 \\ j \neq i}}^{n-1} p_{i,j} a_{j,0} m'_{j,0} = p_{i,0} b_{i,0} + \sum_{\substack{j=1 \\ j \neq i}}^{n-1} p_{i,j} a_{j,0} b_{i,j}$$

得到矩阵形式：

$$(I - Q)(A * M') = (Q * B)A + U * C$$

$$\text{则 } M' = ((I - Q)^{-1}((Q * B)A + U * C)) \div A$$

\bar{m}' 证明略。

证明完毕。

定理 1 给出了在已知实际网络拓扑结构的情况下，对于 overlay 层面的路由算法在网络层面的平均路径长度的计算公式。在第 4 节中，将利用该定理对 2 种路由算法在 overlay 层面和网络层面的性能进行计算分析。

(3) 平均命中率说明

为简化模型，本文暂不考虑节点失效的情况，即不存在进入状态 n 的情况。所以一般情况下平均命中率 $\bar{a} = 1$ 。现分析平均命中率几种特殊情况如下：

1) 非矩阵 U 中元素全为 0， $p_{i,0} = 0 (\forall i \in (0, 1, \dots, n-1))$ ，则 A 一定为 ϕ ， $\bar{a} = \frac{1}{n}$ ；

2) 矩阵 U 某几位不为 0，按照状态转移矩阵元素来表示为 $p_{i,0} \neq 0 (i \in (0, 1, \dots, n-1))$ 。且其他节点到这几个节点的转移概率均为 0，即 $p_{j,i} = 0 (\forall j \in (0, 1, \dots, n-1))$ ，则 A 一定为 ϕ ， $\bar{a} = \frac{1}{n}$ 。

本文中构造转移概率矩阵使得平均命中率 $\bar{a} = 1$ ，从而平均命中率在 overlay 层面和网络层面对平均路径长度的影响一致。下面主要针对平均路径长度的计算来评估结构化 P2P 系统 overlay 层面的路由计算在网络层面的性能。

4 模型模拟与分析

4.1 状态转移概率

本文对平均概率路由和优先级概率路由由 2 种路由算法在

overlay 和网络 2 个层面的性能进行分别比较。下面给出 2 种路由算法对应的转移概率公式(不考虑节点失效的影响)。

假定所有的查找目的地为 0。定义 $l(i)$ 为 i 到目的地 0 的距离, $l(i, j)$ 为 i 到 j 的距离。定义 $N(i) = \{ i' : l(i') = l(i) - 1 \cap l(i, i') = 1 \}$, 可以看出 $N(i)$ 也就是 i 的邻居集合。 $|N(i)|$ 表示 $N(i)$ 集合中的节点个数。以下 $p_{i,j}$ 中的 $i = 1, 2, \dots, n-1$ 。

(1)平均概率路由算法: 在任一节点处都向邻居节点以均等概率转移。则在节点 i 以概率 $\frac{1}{|N(i)|}$ 转移到 $N(i)$ 中节点, 即

$$p_{i,j} = \begin{cases} \frac{1}{|N(i)|}, & j \in N(i) \\ 0, & otherwise \end{cases} \quad (1)$$

(2)优先级概率路由算法: 选路的过程是对邻居表中邻居以一定的优先级排序(以 nodeId 最接近 key 值的邻居优先级最高)来选择下一跳。则在节点 i 转发请求时存在 2 种方式:

1)以概率 1 转移到 $N(i)$ 中优先级最高的节点, 则

$$p_{i,j} = \begin{cases} 1, & j \in N(i) \\ 0, & otherwise \end{cases} \quad (2)$$

其中, j 为 $N(i)$ 中优先级最高的节点。

2)以不等的概率转移到 $N(i)$ 中节点, 原则是优先级高的节点转移概率大。定义 $N(i)$ 中节点为 $v_k (k = 0, 1, \dots, N(i)-1)$, 对应的

的概率为 p_{v_k} , 且 $p_{v_0} > p_{v_1} > \dots > p_{v_{N(i)-1}}$, $\sum_{k=0}^{N(i)-1} p_{v_k} = 1$, 则

$$p_{i,j} = \begin{cases} p_{v_k}, & v_k \in N(i) \\ 0, & otherwise \end{cases} \quad (3)$$

4.2 性能模拟与分析

利用引理和定理 1 计算平均路径长度, 分别计算平均概率路由和优先级概率路由算法在 overlay 和网络两个层面上的平均路径长度, 并进行比较说明。矩阵 Q 、 V 、 B 、 C 是随机产生的符合条件的矩阵。现分别采用式(1)和式(3)来计算平均概率路由和优先级概率路由算法的状态转移概率。需要说明的是: 使用式(3)比式(2)更合理。因为式(2)实际表示的是对某一个 key 值, 从源节点出发往 root 去的路对于途经的每一个中间节点来说, 其下一跳是唯一的。这样在很多情况下存在命中率为 0 的情况。

计算时邻居个数固定为 4, 对系统节点数为 256 和 2048 分别进行 10 次计算, 得到的平均路径长度分别如图 1、图 2 所示。

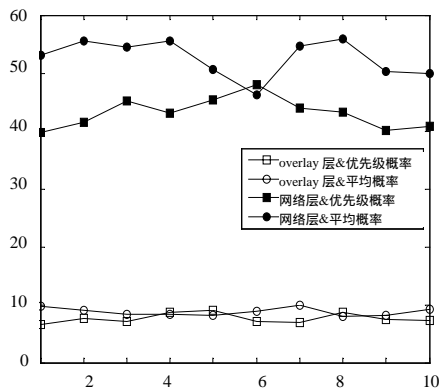


图 1 系统节点数=256 的平均路径长度

由图 1、图 2 可以看出, 在某次计算时, overlay 层面优先级概率分布路由算法计算得到的平均路径长度优于平均概率分布路由算法的计算结果, 而在网络层面中对应点却并不

如此。这说明在 overlay 层面性能优越的路由算法在网络层面并不一定最优。这证实了本文的观点: overlay 层面与网络层面的不匹配导致在 overlay 层面上的距离相近不一定等同于实际网络中的距离相近。按照 overlay 层面计算的路由, 在网络层面上加上实际网络中节点距离的加权后, 可能体现不了该路由的优越性。而实际衡量路由性能往往不是用 overlay 层面路径长度, 而是用包的传输时延, 即对应网络层面的平均路径长度。因此, 在 overlay 层面的路由算法优劣并不能作为实际路由过程的评判标准。

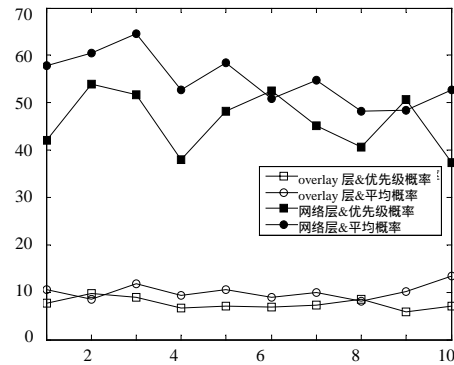


图 2 系统节点数=2048 的平均路径长度

5 总结

本文介绍了通用结构化 P2P 路由系统模型以及 P2P 路由系统马尔可夫链模型。在利用马尔可夫链对结构化 P2P 网络路由系统建模的基础上, 针对 P2P 系统 overlay 层面和网络层面的不一致性, 本文提出了以网络层面平均路径长度为度量的结构化 P2P 网络路由算法在实际网络层性能的量化评估方法。据此进行模拟仿真, 以平均概率路由和优先级概率路由两种路由算法在 overlay 层面和网络层面 2 个层面的性能进行分别比较。分析结果表明, 在 overlay 层面的路由算法优劣并不能作为网络层面也就是实际路由过程的评判标准, 对路由算法的性能评估必须要在网络层面进行。本方法能够准确地表现结构化 P2P 网络 overlay 层面路由算法的实际网络性能, 为结构化 P2P 网络路由算法的研究提供准确的性能评估机制。

参考文献

- 1 Lua E K, Crowcroft J, Pias M, et al. A Survey and Comparison of Peer-to-Peer Overlay Network Schemes[J]. Journal of IEEE Communications Survey and Tutorial, 2005, 7(2).
- 2 Wang Chonggang, Li Bo. Peer-to-Peer Overlay Networks: A Survey[R]. Department of Computer Science, Hong Kong University of Science and Technology, 2003.
- 3 Castro M, Druschel P, Ganesh A, et al. Secure Routing for Structured Peer-to-peer Overlay Networks[C]//Proc. of the 5th Usenix Symp. on Operating Systems Design and Implementation(OSDI). 2002.
- 4 Wallach D S. A Survey of Peer-to-Peer Security Issues[C]//Proc. of the International Symposium on Software Security, Tokyo, Japan. 2002.
- 5 Wang S, Xuan D, Zhao W. On Resilience of Structured Peer-to-Peer Systems[C]//Proc. of IEEE Global Communications Conference (GLOBECOM). 2003.