

P2P 网络节点间链路选取模型的蚁群算法设计与实现

高晓燕

(华北科技学院 计算机学院, 北京 101601)

摘 要: 文中提出了有服务质量保证的虚拟链路的选取模型, 服务网的虚拟链路的选择 QoS 因素主要包括虚拟链路的生存能力和虚拟链路的路由代价等因素, 通过对服务网各参数的分析建立了 P2P 服务网络虚拟链路的选取模型, 并给出了基于模拟退火的蚁群求解算法, 最后通过实验分析了该算法的可行性和有效性。

关键词: 蚁群算法; QoS; 服务; 虚拟链路; 对等网络

中图分类号: TP393.01

文献标识码: A

文章编号: 1000-7180(2013)09-0062-04

Ant Colony Algorithm Design and Implementation for Link Selection Model Between Nodes in P2P Network

GAO Xiao-yan¹

(Department of Computer and Science, North China Institute of Science and Technology, Beijing 101601, China)

Abstract: We put forward a virtual link selection model of service quality assurance, and consider the quality of service metrics in this model includes viability, routing cost and so on. Through the analysis of these network parameters, we establish the virtual link selection model of P2P network and give the ant colony algorithm based on simulated annealing. At last, we analysis the feasibility and effectiveness of the algorithm through the experiment.

Key words: ant colony algorithm; QoS; service; virtual link; P2P

1 引言

虽然目前 P2P 的应用日益广泛, 但仍然缺乏有效的 P2P 拓扑构造机制来保证网络的良性发展^[1]. 尤其新应用的不断出现, 对现有的拓扑模型提出了挑战. 由于现有的 P2P 拓扑模型缺乏对提供有服务质量保证的考虑, 尤其在构建服务网络的虚拟了链路时没有考虑到利用网络中某些节点能够提供强有力的服务质量属性.

构建良好的有服务质量保证的虚拟链路是 P2P 的成功应用的基础, 并且可以提高服务的发现和搜索效率. 据此, 本文构造出一种 P2P 服务网络虚拟链路的选取模型.

2 P2P 服务网虚拟链路选取模型的建立

2.1 P2P 服务网络虚拟链路的 QoS 因数分析

在进行 P2P 服务网虚拟链路选取时, 由于虚拟链路所连接的是端用户, 而端用户的不稳定性必然对虚拟链路的连通性有所影响, 因此在进行虚拟链路的选取时首先要考虑链路的生存能力; 另一方面, 每条虚拟链路都有路由代价, 因此在进行虚拟链路的选取时还要考虑到 P2P 服务网的路由代价问题. 所以, 核心网上虚拟链路的选取设计关键在于使 P2P 服务网的生存能力要高, 并且使 P2P 服务网的路由代价最低. 本文提出的 P2P 服务网的虚拟链路的选择 QoS 因素主要包括以下参数: 虚拟链路的生存能力、虚拟链路的路由代价、虚拟链路的连通性和

收稿日期: 2012-11-20; 修回日期: 2012-12-22

基金项目: 国家自然科学基金项目(61163050); 国家“九七三”计划项目(2011CB311809); 青海省自然科学基金项目(2012-Z-935Q, 2012-ZR-3106)

虚拟链路的延迟。

2.2 P2P 服务网络虚拟链路的形式化描述

由于在进行虚拟链路选取时需要考虑 P2P 服务网的生存能力和连通性,为了给出该模型的形式化描述,首先给出下列分析:

(1)要保证 P2P 服务网的连通性,即 P2P 服务网中任意两个节点 i 和节点 j 之间都是可达的,节点间的连通性不为 $0^{[2]}$,即

$$\gamma_{ij} > 0 \quad (1)$$

(2)P2P 服务网的虚拟链路的生存能力可定义为有效传输的比值。假定给定的节点 i 和节点 j 之间存在多条路径,所有这些路径构成一个集合 $\varphi(i, j) = \{P_k | P_k \text{ 是节点 } i \text{ 和 } j \text{ 之间的第 } k \text{ 条路径}, k=1, 2, 3, \dots, K\}$,其中元素 P_k 是节点 i 和节点 j 之间的第 k 条路径,第 k 条路径的生存概率用 $t(P_k)$ 表示^[3]。网络中传递的信息量用节点间传送的分组集合 $\{d_{ij}\}$ 表示。其中 d_{ij}^k 表示通过第 k 条路径,在 i 及 j 节点之间传送的分组。则 P2P 服务网的虚拟链路节点 i 和节点 j 之间的生存能力 R_{ij} 可描述为:

$$R_{ij} = \max_{P_k \in \varphi(i, j)} \{d_{ij}^k t(P_k)\} \quad (2)$$

(3)在 P2P 服务网模型中, C_{ij} 为 P2P 服务网中虚拟链路节点 i 和节点 j 之间费用代价, C_{ij} 可描述为:

$$C_{ij} = \min_{P_k \in \varphi(i, j)} \{C(P_k)\} \\ C(P_k) = \sum_{P_{k,l} \in \psi(k)} C(P_{k,l}) \quad (3)$$

$\psi(k) = \{P_{k,l} | P_{k,l} \text{ 是节点 } i \text{ 和 } j \text{ 之间第 } k \text{ 条路径的第 } l \text{ 段链路}, l=1, 2, 3, \dots, L\}$, $C(P_{k,l})$ 是节点 i 和 j 之间第 k 条路径的第 l 段链路的费用代价。

(4)在 P2P 服务网模型中, D_{ij} 为 P2P 服务网中虚拟链路节点 i 和节点 j 之间的时延代价,可定义为:

$$D_{ij} = \min_{P_k \in \varphi(i, j)} \{D(P_k)\} \\ D(P_k) = \sum_{P_{k,l} \in \psi(k)} D(P_{k,l}) \quad (4)$$

$\psi(k) = \{P_{k,l} | P_{k,l} \text{ 是节点 } i \text{ 和 } j \text{ 之间第 } k \text{ 条路径的第 } l \text{ 段链路}, l=1, 2, 3, \dots, L\}$, $D(P_{k,l})$ 是节点 i 和 j 之间第 k 条路径第 l 段链路的时延代价。

2.3 P2P 服务网络虚拟链路模型

设 P2P 服务网模型为 $G=(V', E', R)$, 其中 V' 为已选出的节点的集合; E' 为覆盖网中虚拟链路的集合; $R=\{R_{ij} | R_{ij} = \max_{P_k \in \varphi(i, j)} \{d_{ij}^k t(P_k)\}, i \in V', j \in V'\}$, R 为 P2P 服务网的虚拟链路的生存能力,其中元素

R_{ij} 为 E' 的优化链路。 C 为 P2P 服务网的虚拟链路的费用代价, $C=\{C_{ij} | C_{ij} = \min_{P_k \in \varphi(i, j)} \{C(P_k)\}, i \in V', j \in V'\}$; D 为 P2P 服务网中虚拟链路的时延代价, $D=\{D_{ij} | D_{ij} = \min_{P_k \in \varphi(i, j)} \{D(P_k)\}, i \in V', j \in V'\}$ 。在虚拟链路的选取中,假设虚拟链路的费用代价最大不超过 $C_{\max}$,根据以上的分析可得到 P2P 服务网虚拟链路选取问题(PSNLS)的模型可描述为:

$$\max_{P_k \in \varphi(i, j)} \left\{ \frac{d_{ij}^k * I(P_k)}{D(P_k)} \right\}, i \in V', j \in V' \\ \text{s. t. } \begin{cases} \gamma_{ij} > 0 \\ C(P_k) \leq C_{\max} \end{cases} \quad (5)$$

通过求解上式,根据 P2P 服务网上虚拟链路的生存能力与 P2P 服务网代价进行虚拟链路的选取,一方面能使虚拟链路的延迟最小,另一方面使得 P2P 虚拟链路的生存能力最大,从而可以减少节点维护的 P2P 服务网状态信息,利于 P2P 服务网的稳定性。基于上述讨论,本文采用基于模拟退火的蚁群算法来解决 PSNLS 模型。

3 求解 PSNLS 问题的算法

从公式(8)可以看出,PSNLS 问题是典型的组合优化问题,是一个 NPC 问题^[3],当 P2P 服务网规模比较大时,问题的时间复杂度很大,用一般的方法不能解决这个问题,因此在这种情况下可以考虑运用蚁群算法来求解。

(1) 蚁群算法基本原理

蚁群算法的特点是它是一个增强型学习系统,具有分布式的计算特性,具有很强的鲁棒性。蚂蚁在觅食的过程中在路径上释放出一种特殊的信息素,后面的蚂蚁根据遗留下的信息素选择下一步要走的路径。路径上的信息素值越高,蚂蚁选择这条路径的概率就越大,构成一个学习信息的正反馈过程。蚁群算法正是在对蚁群行为的研究基础上提出的一种启发式算法^[4-5]。

蚁群算法可以应用在 P2P 服务虚拟链路的选取方面。此时,在蚁群算法中,每个蚂蚁从 P2P 服务网中一个节点出发,根据状态转移规则选择下一跳节点,直到此蚂蚁走完所有节点。蚂蚁所走的路径是根据路径上的信息素的数量来进行选取的。

当 m 只蚂蚁成功地完成一次寻径行为后,选择出目标函数值最大的链路,来进行全局信息素更新。若 i, j 是两个相邻的节点,其中 $\tau_{ij}(t)$ 为时间 t 蚂蚁留在路径 (i, j) 上的信息量,每只蚂蚁在时间 t 开始

一次新的循环,每次循环蚂蚁为所有目标节点都选择一次路由.一次循环结束时间更新为 $t+n$, 蚂蚁会根据下式更新路由路径 (i,j) 上的信息量:

$$\tau_{ij}(t+n) \leftarrow (1-\rho)\tau_{ij}(t) + \rho \times \Delta\tau_{ij}(t)$$

式中 $0 < \rho < 1$ 为常数, $1-\rho$ 为 $\tau_{ij}(t)$ 在时间 t 和 $t+n$ 之间的挥发程度.

$$\Delta\tau_{ij}(t) = \sum_{k=1}^m \Delta\tau_{ij}^k(t)$$

m 为蚂蚁总数, $\Delta\tau_{ij}^k(t)$ 为在 t 和 $t+n$ 之间, 由第 k 只蚂蚁引起的路由路径 (i,j) 上信息量的变化.

$$\Delta\tau_{ij}^k(t) = \begin{cases} Q/L_k, & \text{第 } k \text{ 只蚂蚁选择路由路径 } (i,j) \\ 0, & \text{其他} \end{cases}$$

其中 Q 为常数, L_k 为蚂蚁 k 求得的 P2P 服务网路由代价总和. 蚂蚁 k 以概率 $P_{ij}^k(t)$ 选择路由路径 (i,j) 作为本次循环路径.

$$P_{ij}^k(t) = \frac{(\tau_{ij}(t))^\alpha (\eta_{ij})^\beta}{\sum_{j' \in \text{neighbor}(i)} (\tau_{ij'}(t))^\alpha (\eta_{ij'})^\beta}$$

其中, η_{ij} 为启发式信息, $\eta_{ij} = \frac{d_{ij}^{P_k} * t(P_k)}{D(P_k)}$, α, β

为常数, 分别表示 $\tau_{ij}(t)$ 和 η_{ij} 的重要程度. 这样每只蚂蚁为所有的目标节点都选择了一个路由路径, 将所有目标节点的路由路径进行综合, 然后去掉重合的边, 就得到从一个节点出发到所有其它所有节点 P2P 服务网.

(2) 求解 PSNLS 问题的算法—基于模拟退火的蚁群算法 SANT

为了提供蚁群算法的全局搜索能力和搜索速度, 可以对传统蚁群算法进行如下改进:

① 保留最优解, 在每次循环结束后, 求出最优解, 将其保留.

② 在算法运行的初始阶段, 每个路由路径上的信息量相差不大, 通过信息的正反馈, 较好解的信息量增大, 从而逐渐收敛. 当问题规模比较大时, 此过程需要较长的时间. 为提高算法的收敛速度, 引入模拟退火算法^[6-7], 在每次循环结束, 从蚁群算法搜索过程中得到的中间解将被作为模拟退火算法的初始解. 在模拟退火过程中, 解的接受概率将服从 Bolt-zmann 概率分布.

算法具体步骤描述如下:

Begin

Initialization the network parameter

$$t=0, \tau_{ij}(t)=0, \Delta\tau_{ij}(t)=0, \eta_{ij} = \frac{d_{ij}^{P_k} * t(P_k)}{D(P_k)};$$

While (not termination condition)

{For ($k=1; k \leq m-1; k++$)

{Locate m ants in the node r , which is selected randomly}

For ($i=1; i \leq n; i++$)

{For ($k=1; k \leq m-1; k++$)

{ The ant ak selects routing path for node i with a probability p_{ij}^k }

For($k=1; k \leq m; k++$)

{

If the result meet request, then Compute the best result and give it the m ant.

Else CALL SA()

Compute the network dependability γ_0 .

If $\gamma_{ij} > \gamma_0$

{Compute the network topology aANTin}

Put out the optimization result.

End }

Algorithm SA()

$$f(x) = \min_{P_k \in \varphi(i,j)} \left\{ \frac{d_{ij}^{P_k} * t(P_k)}{D(P_k)} \right\}, T = ST_{\max},$$

$x=x_t$, x_t is the first result in the FpstackvS

FOR ($T=ST_{\max}; T < ST_0; T--$)

{Computer $U(x_t), x' = U(x_t)$;

IF

$\Delta f = f(x') - f(x) < 0, x = x'$

Else

$$\exp\left(-\frac{E(x_t) - E(x)}{t}\right) > \text{random}(0,1)$$

则令 $x = x_t$;

Put out the optimization result}

Return();

END

SANT 算法中计算出每个节点到其他节点的最大生存能力路径, 并使此时 P2P 服务网连通性满足约束条件, 最后合并所有 P2P 服务网拓扑结构, 删除重复路径, 从而构成一个最终的 P2P 服务网. 此 P2P 服务网在满足 P2P 服务网可靠性约束条件的基础上减少 P2P 服务网的虚拟链路的数量, 从而为在 P2P 服务网上提供更好的服务性能^[8-9].

(3) SANT 算法分析

为了验证本文 SANT 算法的性能, 我们同时对蚁群算法(ANT)和此算法进行了仿真模拟, 通过与蚁群算法进行比较来证明 SANT 算法的正确性和

具有更好的收敛性. 仿真网络采用基于 Waxman 模型的拓扑生成算法^[9-10]. $\alpha=0.15, \beta=0.2$; 其中 P2P 网络的节点占全部网络节点数的 20%, 均匀分布在网络中; 每个节点所连接的端用户数量权值 $w(i)$ 在 $[0, 5]$ 上均匀分布. 网络节点的数目从 10 逐渐增加到 50 个, 在不同约束条件下进行比较实验, 仿真结果如图 1 和图 2 所示.

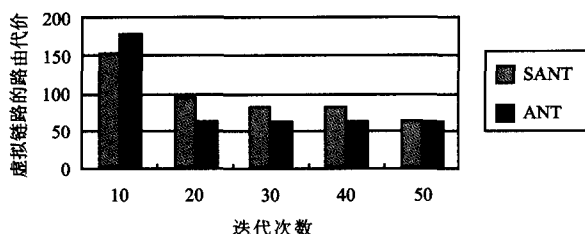


图1 两种算法收敛速度比较

图1是网络节点数为30时, SANT和ANT算法求得P2P网在满足约束条件的路由代价所计算的迭代次数. 从图中可以看出, 在相同网络规模条件下, SANT算法具有更快的收敛速度, 比ANT算法具有更好的性能.

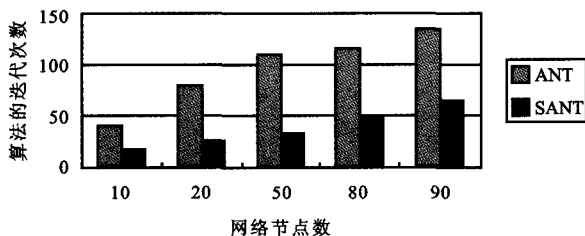


图2 网络节点变化时迭代次数

图2为两种算法在不同网络规模条件下的收敛速度, 从图中可以看出, 随着网络节点的增多, SANT算法具有更快收敛的速度, 这是因为随着网络规模的增加, 算法的可行解的空间增多, 如果进化算法不能够有效的缩小解空间, 必然不能提高算法的收敛速度. 而SANT算法利用模拟退火的思想, 从而保证可行解的收敛, 最终加快迭代计算找到最优解(或近似最优解)的速度.

4 结束语

本文根据P2P服务网络的特点, 建立了能保证服务质量的P2P网络虚拟链路的数学模型, 然后设计了该模型的求解算法, 通过实验分析了蚁群算法

的可行性和有效性. 下一步工作是完善实验, 并通过实验的反馈来改进建立的数学模型, 使模型更好的反映P2P网络的链路的实际情况, 有效地提高网络的服务发现速度.

参考文献:

- [1] Kant K, R. Iyer, Tewari V. A framework for classifying peer-to-peer technologies[C]//Proceedings of the 2nd IEEE/ACM International Symposium on Cluster Computing and the Grid (CCGRID'02). Germany, Berlin, 2002(15):200-222.
- [2] Aberer K, Hauswirth M. An overview on peer-to-peer information systems[C]//Proceedings of Workshop on Distributed Data and Structures (WDAS-2002). Paris, France, 2002(11): 100-112.
- [3] 卢开澄. 单目标、多目标与整数规划[M]. 北京:清华大学出版社, 1999.
- [4] Dorigo M, Birattari M, Stützle G. Ant colony optimization[J]. IEEE Comput. Intell. Mag., 2006, 1(4): 28-39.
- [5] Marco Dorigo, Christian Blum. Ant colony optimization theory: a survey[J]. Theoretical Computer Science, 2005, 344(1/2): 243-278.
- [6] Gutjahr W J. Mathematical runtime analysis of ACO algorithms: survey on an emerging issue[J]. Swarm Intell., 2007, 1(1): 59-79.
- [7] Nattapat Attiratanasunthron, Jittat Fakcharoenphol. A running time analysis of an ant colony optimization algorithm for shortest paths in directed acyclic graphs[J]. Information Processing Letters, 2008, 105(3): 88-92.
- [9] 姚新, 陈国良, 徐惠敏, 等. 进化算法研究进展[J]. 计算机学报, 1995, 18(9): 694-705.
- [10] Stu Tzl E T, Dorigo M. A short convergence proof for a class of ant colony - optimization - algorithms[J]. IEEE Transactions on evolutionary computation, 2002, 6(4): 358-365.

作者简介:

高晓燕 女, (1970-), 博士, 副教授. 研究方向为计算机网络、网络安全.