维普资讯 http://www.cqvip.com

适用于内容分发网络的动态负载均衡策略*)

徐卫东[1] 王 康[2]

(东软金算盘软件有限公司 重庆 400020)[1] (重庆大学计算机科学与工程学院 重庆 400044)[2]

摘 要 本文阐述了内容分发网络和网络负载均衡技术的基本原理、算法,分析了现有负载均衡算法及BP神经网络的优缺点,在此基础上,将神经网络理论应用于网络负载均衡,提出了基于改进BP算法并适用于内容分发网络的负载均衡算法和实现方法。

关键词 内容分发网络,负载均衡,人工神经网络,BP算法

Dynamic Load-Balancing Strategy for Content Delivery Network

XU Wei-Dong[1] WANG Kang[2]

(Neusoft Gold Abacus software Ltd.,Chongqing 400020)[1]

(Department of Computer science,Chongqing University,Chongqing 400044)[2]

Abstract In the article the theory of Content Delivery Network and Load-Balancing has been introduced. At the same time some limitations of Load-Balancing algorithm and BP neural network have been analyzed,based on that, theory of neural network is implied in the Load-Balancing of the network,Load-Balancing strategy for Content Delivery Network and its merits based on Improved BP algorithm have been put forward.

Keywords Content delivery network,Load-Balancing,Artificial neural network,BP algorithm

1 引言

随着Internet的迅速普及和用户数量的急剧增加,人们对Internet内容和应用的需求正在飞速增长,因此服务器需要具备提供处理大量并发访问的能力,而服务器的处理和I/O能力成为了提供服务的瓶颈,使得越来越多的Web站点发现他们很难提供给用户满意的网络服务质量。尽管,目前网络带宽已成倍增长,但带宽并不是影响互联网连接速度快慢的唯一因素,如果没有良好的负载均衡策略,网站不管有多大带宽,用户对网站的访问连接同样会非常慢。而要很好地解决这些问题,必须要有新技术的支持,内容分发网络技术因此应运而生,它是一套智能的解决方案,其实质是通过为网络内容或服务建立多个备份并采用多种机制将其复制到特定位置,然后把用户访问根据某种负载均衡策略重定向到最近或网络状况最佳的原始服务器的内容备份节点,实现网络内容就近访问,以减少访问时延提高网络资源利用率。

内容分发网络的核心技术主要包括负载均衡、内容分发与复制、内容路由等,其中负载均衡技术对于各类计算机网络系统来说都具有非常重要和普遍的意义,因此,本文重点探讨内容分发网络的负载均衡技术。

2 内容分发网络的基本原理

内容分发网络(Content Delivery Network 以下简称CDN)是建立在互联网之上的一层特殊网络,专门用于通过互联网高效传递丰富的多媒体内容。其网络架构主要由两大部分组成,即中心和边缘两部分,中心指CDN网管中心和全局负载均衡DNS,负责全局负载均衡和内容路由,边缘主要指异地内容备份节点,是CDN分发的载体。

当用户访问加入CDN后,域名解析请求将最终交给CDN全局负载均衡DNS处理。负载均衡DNS通过定义好的策略,将用户请求重定向到当时最接近用户或网络状况最佳的备份节点,使用户能够得到快速服务,并减轻了原服务器负载,均衡了网络流量,最终消除网络拥塞。同时,它还与分布在各地的所有CDN备份节点保持通信,搜集各节点的通信及负载状态,确保不将用户请求分配到不可用的节点,实际是通过DNS做全局负载均衡。

3 网络负载均衡基本原理及算法

用户访问Internet时,在多个拥有同样内容的服务器中,总喜欢根据经验或监测数据访问熟悉或负载较小的服务器。CDN系统则可以通过定期监测服务器状态并将请求指向最佳服务器来实现请求的动态分配。这种在多个服务器中根据服务器负载动态定向请求的技术称为动态负载均衡。CDN中该功能可以加入DNS中,因为DNS本身就充当了解析客户请求的责任,具有这种特性的DNS称为动态负载均衡DNS。

动态负载均衡系统考虑的关键问题是及时、准确地把握节点负载状况及网络带宽,并根据各节点当前负载状态动态调整任务分布。一般集群节点可提供系统负载、用户数、流量及其他资源信息,并通过高效通信机制传给负载均衡器,均衡器监视所有节点状态,决定下个任务传给谁。均衡器可以是单个设备,也可以是一组平行或树状层次分布的设备。

基本负载均衡算法如下:

(1)*轮转法* 是最简单的调度算法。在一个任务队列里,轮转法简单地在这组成员节点中顺序轮转选择。它适用于各节点处理能力和性能均相同的情况。而对不同性能特点的忽视是该算法的缺陷。

(2)*最少连接法* 均衡器记录当前活跃连接,把下一个请求发给含有最少连接数的节点。但由于不同应用对系统资源消耗的差异很大,连接数无法反映真实负载,为减少此不利影响,可对各节点设置最大连接数上限阈值。

*)基金项目:中国科学院知识创新工程项目"IPv6网络关键技术研究与城域示范系统";重庆市科技计划项目(2003044)。徐卫东 硕士,主要研究方向:计算机网络;王 康 教授,主要研究方向:信息网络安全技术、Internet网络技术。

维普资讯 http://www.cqvip.com

(3)*最快响应法* 均衡器记录到每一个集群节点的网络响应时间,并将下一个连接请求分配给响应时间最短的节点,该方法要求使用 ICMP 包或基于 UDP 包的专用技术来探测各节点。该算法不太适用于 LAN,因为 LAN 中各节点响应时间差异极小;而在 WAN 上,响应时间对用户就近选择服务器而言就具有现实意义;且集群拓扑越分散该方法效果越好。

(4)*加权法* 只能与其他方法合用,起很好的补充作用。该算法根据节点优先级或当前负载状况、响应时间等(权值)来构成负载均衡的多优先级队列。这里权值是基于各节点能力的估计值。

(5)*排名算法* 基于如下用户数目和负载平均的列表。此算法是较合理的,因为它根据最少单个访问以及较低负载平均来进行排名最佳主机。该算法在动态负载均衡 DNS 中确定最佳服务器时被使用。

```
WT_PER_USER = 100
USER_PER_LOAD_UNIT = 3
FUDGE = (TOT_USER - UNIQ_USER) * (WT_PER_USER/
        5)
WEIGHT = (UNIQ_USER * WT_PER_USER) + (USER_
PER_LOAD_UNIT * LOAD) + FUDGE
```

变量名称含义如下:

TOT_USER:登录用户总数

UNIQ_USERS:登录的不重复用户数目

LOAD:最后一分钟的负载平均乘 100

WT_PER_USER:每个用户负载量

FUDGE:用户多次登录后的修正参数

WEIGHT:服务器排名

4 基于改进 BP 算法的动态反馈负载均衡策略

4.1 神经网络理论及其典型网络模型

人工神经网络是由大量简单神经元相互联结,模拟人脑神经处理信息的方式,进行信息并行处理和非线形转换的复杂网络系统。它具有极强的非线形逼近、大规模并行处理、自训练学习、自组织和容错能力等优点。

按照网络拓扑结构和运行方式,神经网络模型分为前馈多层式网络模型、反馈递归式网络模型、随机型网络模型等。目前应用最为广泛的是前馈多层式网络中的 BP 反向传播模型。

4.2 BP 神经网络基本原理

BP 模型是一种用于前向多层神经网络的反传学习算法。BP 算法用于多层网络,网络中有输入层节点及输出层节点和一层至多层隐层节点,如图 1 所示。当信息输入网络时,要先向前传播到隐含层节点,经作用函数后,再把隐节点的输出信号传播到输出节点,最后给出输出结果。节点激励函数通常选取 S 型函数,如:

$$f(x)=1/1+e^{-x/Q}$$

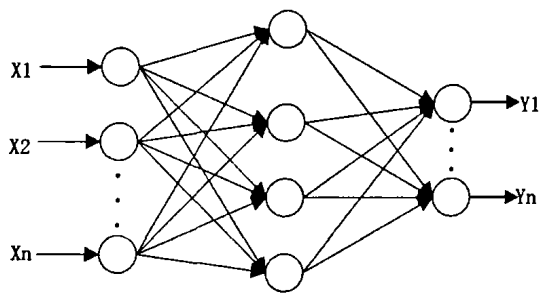

图 1 BP 网络

上式中 Q 为调整激励函数形式的 Sigmoid 参数。该算法的学习过程由正向传播和反向传播组成。输入信息从输入层经隐含层逐层处理,并传向输出层。每一层神经元状态只影响下一层神经元状态。如果输出层得不到期望的输出,则转入反向传播,将误差信号沿原来的连接通道返回,通过修改各层神经元的权值,使得误差最小。

4.3 BP 网络模型的改进策略

(1)*学习因子 η 的优化* 采用变步长法根据输出误差大小自动调整学习因子,来减少迭代次数和加快收敛速度。

$$\eta=\eta+a(Ep(n)-Ep(n-1))/Ep(n) \quad (1)$$

a—调整步长,0~1 间取值

(2)*隐层节点数的优化* 隐节点数的多少对网络性能影响较大,当隐节点数太多时,会导致网络学习时间过长,甚至不能收敛;而当隐节点数过小时,网络容错能力差。利用逐步回归分析法并进行参数的显著性检验来动态删除一些线形相关的隐节点。当由该节点出发指向下一层节点的所有权值和阈值均落于死区(一般取±0.1、±0.05 等区间)中,则可删除。最佳隐节点数 L 可参考下面公式(2)计算,或中间层神经元数取 $2m-1$。

$$L=(m+n)1/2+c \quad (2)$$

m—输入节点数;n—输出节点数;c—介于 1~10 的常数。

(3)*权调节向量* BP 算法是基于最速下降法的,由于梯度最速下降法的固有缺点易陷入局部极小、收敛速度慢和引起振荡效应,Rumelhart,Hinton 和 Williams(1986)建议在权调节向量中增加"惯性量"。即:

$$\Delta\omega_{ij}(t+1)=\eta\delta_{j}o_{i}+\alpha\Delta\omega_{ij}(t) \quad (3)$$

式中:$\omega_{ij}(t+1)$,$\omega_{ij}(t)$分别表示第 $t+1$ 次、第 t 次迭代的权值修正量;α 为动量因子 $\alpha\in(0,1)$,η 为学习速率因子;$\delta_{j}o_{i}$ 表示 BP 算法中的误差平方和对权值的负梯度。在训练中先取较小的 α 值,当结果接近期望值时适当加大 α(或同时加大 η 值)进行学习训练,直到满足要求为止。本文采用了该方法进一步加速了收敛速度,并在一定程度上减少了陷入局部极小的概率。

4.4 改进 BP 算法在 CDN 负载均衡算法中的应用

基于 BP 算法的动态反馈负载均衡是对 3 节中算法加以综合改进,并考虑各节点的实时负载和响应能力,利用 BP 网络预测链路状态并转换为节点权值,根据权值来不断调整任务分布比例,避免有些结点超载时仍收到大量请求,从而提高 CDN 整体吞吐率。

在 CDN 内,负载均衡 DNS 上运行服务端监控进程,监控进程负责监视和收集 CDN 各备份节点负载相关状态信息;而每个节点上运行代理进程,负责定时向均衡器报告自身负载状况。监控进程根据收到的全部节点的负载信息来进行同步操作,即对将要分配的任务按照权值的比例重新分布。新权值计算主要根据各节点 CPU 利用率、可用内存、磁盘 I/O 状况,带宽等计算,若新权值和当前权值的差值大于设定阀值,监控器采用新权值对 CDN 内的任务重新分布,直到下一次负载信息同步到来之前。均衡器可以配合动态权值,采用加权轮询算法来对接受的网络服务请求进行调度。

应用实现的基本思想是应用神经网络方法建立起备份节点负载状况特征参数与备份节点和实际请求发出节点之间链路状态(可以是传输时延、传输速率等)的非线性映射关系模型。采用改进 BP 算法来训练前馈神经网络,即将 CDN 中负载均衡 DNS 从各备份节点定时采集到的运行时负载状况特征参数和节点与请求发出节点间实际测试到的链路时延经转换后作为神经网络的训练样本,这里请求发出节点是指负载均衡 DNS 分配给备份节点的连接请求的发出主机。而实际链路时延在备份节点与请求节点建立连接后才能测到,具有滞

维普资讯 http://www.cqvip.com

后性，因此，将节点负载状况特征参数作为 BP 网络输入，预测链路传输时延作为输出，然后将预测到的传输时延通过模式转换为节点权值。将训练好的网络用于预测传输时延并转换为节点权值，负载均衡系统就能很容易地根据节点权值进行网络流量动态分配，从而实现 CDN 的负载均衡，使 CDN 负载状况达到近似理想状况。

此外，关于采集的周期，虽然很短的周期可以更确切地反映各个节点的负载，但是频繁地采集会给均衡器和节点带来负担和增加网络开销。此外，由于采集时刻进行负载计算，均衡器反映出来各节点的负载信息会出现剧烈抖动，均衡器无法准确反应备份节点真实的负载变化趋势。因此应适当调整采集负载信息的周期，如设间隔 $t=300s$。

4.4.1 *基于改进 BP 神经网络的动态反馈负载均衡模型*（如图 2 所示）

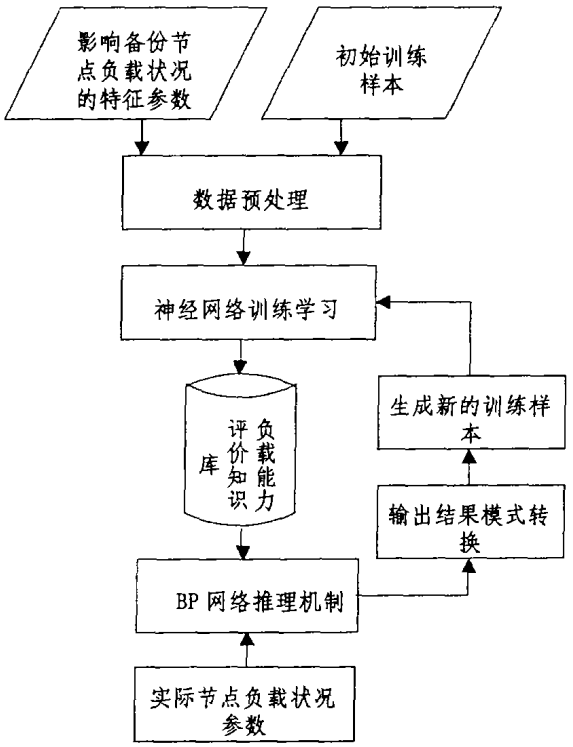

图 2 基于改进 BP 神经网络的动态反馈负载均衡模型

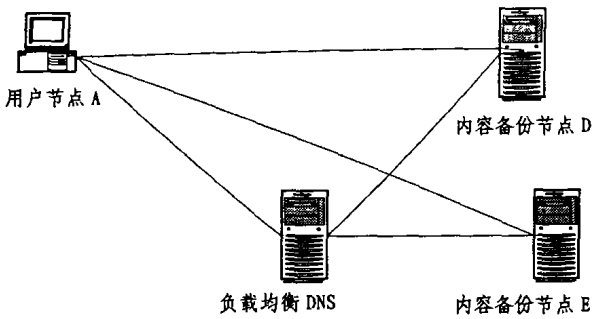

图 3 CDN 负载均衡示意图

4.4.2 *负载均衡的实现* 本文以图 3 的简单网络结构来讨论负载均衡策略的实现。设所有内容备份节点都拥有原始服务器相同的内容备份，即为 CDN 用户节点 A 分配最佳内容备份服务节点。

(1)算法实现的步骤

① 用户节点 A 发出服务请求经域名解析到达 CDN 负载均衡 DNS，DNS 通过提取 A 的 IP 地址，得出 A 的网络号 N。

②负载均衡 DNS 搜索保存的 CDN 路由信息表如表 1 所示，先排出不可达的节点（一般 DNS 认为采集周期内未报告负载信息的备份节点为不可达节点，或者由备份节点主动发出拒绝服务报文）。在可用节点中如果发现仅有一个备份节点网络号与 N 相同即与 A 在同一个网内，则直接将 A 的请求重定向至该节点；如果发现多个备份节点网络号与 N 相同，且没有与 N 相同网络号的下层负载均衡 DNS，则采用加权排名算法进行任务分配；如果发现多个备份节点网络号与 N 相同，且有与 N 相同网络号的下层负载均衡 DNS，则将请求重定向到该下层负载均衡 DNS；如果发现没有备份节点网络号与 N 相同，则采有改进 BP 算法计算所有备份节点权值，并用计算所得权值作为加权轮询算法权值的依据。

表 1 CDN 路由表示意图

目的 URL 地址	备份节点地址
www.content1.com	10.0.0.1
www.content1.com	20.0.0.1
www.content1.com	30.0.0.1
www.content2.com	10.0.0.2
www.content2.com	10.0.0.3

(2)加权轮询算法。该算法用节点权值表示节点的处理性能。并根据权值的高低顺序按照轮询的方式将任务请求分配到各节点。加权轮询的基本原理可描述为：

假设某集群内有一组节点 $N=\{N_0,N_1,\cdots,N_{n-1}\}$，$W(N_i)$ 表示节点 N_i 的权值，指示变量 $i\in\{0,1,2,\cdots,n-1\}$，$T(N_i)$ 表示节点 N_i 当前所分配的任务量。

$\Sigma T(N_i)$ 表示当前同步周期需要处理的任务总量。

$\Sigma W(N_i)$ 表示节点权值总和。

则 $$W(N_i)/\Sigma W(N_i)=T(N_i)/\Sigma T(N_i) \quad (4)$$

表示任务分配是按照各节点权值占权值总数的比例来进行的。即将请求分配给当前负载状况最好的内容备份节点。

(3)用改进 BP 算法预测链路传输时延并转换为节点权值。当备份节点初次加入系统时，系统管理员根据结点的硬件配置情况对每个节点都设定一个初始权值，在负载均衡器上也先使用这个权值。然后，随着节点负载的变化，对权值进行调整。

① 确定特征参数和状态参数，网络的拓扑结构，包括中间隐层的层数，输入层、输出层和隐层的节点数：

文[7]中已证明，含一个隐含层的 BP 网络能够实现任意非线性物理关系的连续映射，因此本文用三层 BP 网络建立备份节点负载状况特征参数与节点和实际请求发出节点之间链路状态（可以是传输时延、传输速率等）的非线性映射关系模型。

文中仅选取了对负载影响最大的几个因素如：CPU 资源、I/O 状态、内存资源、静态内容服务连接数、动态内容服务连接数、带宽、节点最大用户数、响应时间作为传输时延预测的特征参数，形成 BP 网络的 8 个输入层节点。选取备份节点到实际请求发出节点的链路传输时延预测值为 BP 网络的单输出，实际链路传输时延由备份节点在下一个信息采集周期传回。根据 4.3 节策略可确定初始学习因子 $\eta=0.9$，初始动量因子 $a=0.5$，预设误差为 0.0001，单隐层，其隐节点数为 $L=15$，构成 8-15-1 *BP* 神经网络模型。在系统训练时，实际还要对不同的隐层节点数分别进行比较，最后确定出最合理的网络结构。

② 初始权值的确定：初始权值应是不完全相等的一组值。故而，应设计一个随机发生器程序，产生一组 $-0.5\sim+0.5$ 的随机数，作为网络的初始权值。

维普资讯 http://www.cqvip.com

③ 迭代次数：一般取1000次以上。由于神经网络计算并不能保证在各种参数配置下迭代结果收敛，当迭代结果不收敛时，允许最大的迭代次数。

④ 数据转换：可以对输入层各个节点的数据进行转换，提供转换的方法有取对数、平方根转换和数据标准化。下面给出归一化公式(5)，使计算过程中Sigmoid函数不至于溢出：

$$x' = 0.8x - x_{min}/x_{max} - x_{min} + 0.1 \quad (5)$$

式中：x 代表归一化前训练模式的各分量值；x_{max}、x_{min} 分别表示训练模式分量值中的最大者和最小者；x' 代表归一化后训练模式的分量。

归一化的目的是使训练模式分量值保持在区间[0.1，0.9]之内，使网络计算时在网络规模有限的情况下不会溢出。这种线性变换的归一化并未使模式产生性质上的变形，从而保证信息没有失真。

⑤ 选择学习样本：选取多组对应不同权值时的特征参数值作为学习样本，供网络系统学习。这些样本应尽可能地反映节点各种负载状态。神经网络的学习过程即根据样本确定网络的联接权值和误差反复修正的过程。

⑥ 确定作用函数，通常选择非线形S型函数，Sigmoid调整神经元激励函数形式，一般取0.9～1.0之间。

⑦ 建立系统负载能力评价知识库。通过网络学习确认并保存的网络结构包括：输入、输出和隐节点数以及反映其间关联度的网络权值、阈值的组合；即为具有推理机制的负载评价知识库。

⑧ 实际系统中应用。将实际备份节点负载状况特征值转换后输入到已经过训练具有推理功能的BP网络中，运用系统负载能力评价知识库处理后得到请求节点与备份节点链路状态预测值并转换为节点权值。实际系统测试结果又作为新的学习样本输入神经网络，使知识库进一步充实。当实际值与预测值差超过阈值时对BP网络重复训练。

结论 基于改进BP算法的动态反馈负载均衡策略去除了复杂的管理员人工调节，而动态地根据服务器负载和性能来调节。而且在传统加权轮询算法中，权值仅由服务节点负载状况决定，并没有考虑实际请求发出节点与备份节点间的链路状态，但在CDN系统中用户节点与内容备份节点间的链路状态又是非常重要的。而基于BP算法的动态负载均衡算法同时考虑了以上两方面因素以及服务请求的不同类型、后台服务器的不同能力，利用神经网络良好的自学习、自适应、非线形逼近能力预测链路状态，并通过将链路状态的预测值转换为权值进行用户请求的动态分配，因此能很好地适应CDN负载均衡的要求。

参考文献

1 谢希仁. 计算机网络[M]. 北京：电子工业出版社，1999
2 戴葵. 神经网络实现技术[M]. 北京：国防科技大学出版社，1998
3 王永庆. 人工智能原理与方法[M]. 西安：西安交通大学出版社，1997
4 从爽，赵何. 反向转播网络的不足与改进[J]. 自动化博览. 1999，1：25～26
5 Stevens W R. TCP/IP Illustracted Volume 1：The Protocols [M]. 北京：机械工业出版社，2000
6 Nilsson N J. Artificial Intelligence A New Synthesis [M]. 北京：机械工业出版社，2000
7 Cybenko D H. Approximation by superpositions of a sigmoidal function [J]. Math. Control，Signals，Systems，1989，2：303～314
8 Tanenbaum A S，Steen M v. Globule：an Open-Source Content Distribution Network [EB/OL]. http://www. globule. org/，2000-2004

（上接第33页）

·SAMax/MetroMax/温度更新系数与运算时间的关系。从图3可以看出，运算时间随着参数SAMax及MetroMax的增加而明显增大，而温度更新系数的取值对于运算时间的影响不是特别明显。

同样，对EDFAL、PL以及费用参数等也做了类似的评价，详见文[10]。

5.2 基于SGA的智能算法参数评价及其同基于SAA的比较

对SGA涉及到和参数及其仿真结果进行了评价并且与基于SAA的智能算法进行了对比分析[10]。

表1是对两种算法在相同条件下建立连接数、配置WC数和运行时间的比较。比较时，以SGA所得结果为基准。从表1可以看出，SGA性能总体上优于SAA。

表1 WC配置结果比较

SGA		SAA
建立连接数之比 (SAA：SGA)	配置WC数之比 (SAA：SGA)	运算时间之比 (SSA：SGA)
0.85：1	1.4：1	0.44：1
1.02：1	1.2：1	0.5：1
0.91：1	0.83：1	0.43：1
0.98：1	1.6：1	0.58：1
0.99：1	1.17：1	0.5：1
0.93：1	0.75：1	0.45：1
1.02：1	1.14：1	0.48：1

结束语 本文在保证网络性能的前提下，综合考虑了网络拓扑、波长资源、波长转换器、波长放大器以及功率等多种约束条件，对WA问题进行求解。采用分层解决方案，运用软计算与启发式相结合的方法来具体实现，尽可能满足用户连接请求。此外，设计并实现了一个仿真软件。仿真结果表明，算法是可行和有效的。

参考文献

1 Al-Salameh D Y，Fatehi M T，Gartner W J，et al. Optical networking. Bell Labs Technical Journal，1998. 01，3(1)：39～61
2 Shen Gangxiang，Bose S K，Cheng Tee Hiang，et al. Efficient wavelength assignment algorithms for light paths in WDM optical networks with/without wavelength conversion. Photonic Network Communications，2000. 06，2(4)：349～359
3 Ye Yabin，Zhang Hanyi，Qin Tao，et al. Statistics study of routing and wavelength assignment algorithms in WDM all optical network. Optics Communications，2000. 11，185(4-6)：315～320
4 Podcameni A，Lopes J. Using a simple algorithm and platform in optical DWDM networks for reaching a satisfactory wavelength-routing assignment. Microwave and Optical Technology Letters，2001. 03，28(6)：406～410
5 Evripidis Bampis，Rouskas George N. The scheduling and wavelength assignment problem in optical WDM networks. Journal of Lightwave Technology，2002. 05，20(5)：782～789
6 Ozdaglar Asuman E，Bertsekas Dimitri P. Routing and wavelength assignment in optical networks. IEEE/ACM Transactions on Networking，2003. 04，11(2)：259～272
7 Josue K，Nicolas P，Maurice G，et al. Routing and wavelength assignment of scheduled lightpath demands. IEEE Journal on Selected Areas in Communications，2003，21(8)：1231～1240
8 王兴伟，程辉，李佳，黄敏. 一种IP/DWDM光因特网中的组播路由算法. 东北大学学报(自然科学版)，2003，24(12)：1165～1168
9 Maher A，Byrav R，Deogun Jitender S. Routing and wavelength assignment with power considerations in optical networks. Computer Networks，2000，32(5)：539～555
10 刘聪. IP/DWDM光因特网中多约束波长分配算法的仿真研究与实现[学位论文]. 沈阳：东北大学，2003
11 刑文顺，谢金星. 现代优化计算方法. 北京：清华大学出版社，1999
12 Chu Xiaowen，Li Bo，Chlamtac Imrich. Wavelength converter placement under different RWA algorithms in wavelength-routed all-optical networks. IEEE Transactions on Communications，2003，51(4)：607～617