

文章编号:1001-9081(2006)07-1563-04

基于快速传输协议实现卫星 TCP 性能的改善

吴 结,高随祥

(中国科学院 研究生院,北京 100049)

(jonathan@publicb.bta.net.cn)

摘 要:在经由卫星通信系统的网络应用中, TCP 的性能较低。分段连接机制作为应用最广的一种性能优化方案,可以改善 TCP 端对端的性能,而在此机制中,卫星链路部分采用何种传输协议直接决定性能改善的效果。XTP 协议是一种应用于有线网络的高速传输协议,但经过深入研究发现: XTP 协议同样也非常适用于卫星通信系统。所以建议:在分段连接机制中,采用 XTP 协议作为卫星链路的传输协议。性能测试结果表明采用该建议非常显著地提高了端对端 TCP 的性能。

关键词:TCP 性能;分段连接机制;快速传输协议;错误控制机制;速率控制机制

中图分类号: TP393.03 **文献标识码:** A

Performance improvement of TCP over satellite link based on the express transport protocol

WU Jie, GAO Sui-xiang

(Graduate School, Chinese Academy of Sciences, Beijing 100049, China)

Abstract: In the network applications over satellite communication system, the TCP performance over satellite link is poor. The splitting TCP mechanism, which is widely applied, can improve the end-to-end performance of TCP. How to choose transport protocol for satellite link determines the performance of whole mechanism. The eXpress Transport Protocol (XTP), adopted in wire network, is a high speed transport protocol. After deep research for the protocol, it was found that XTP could be applicable to satellite communication. So XTP was proposed to be adopted as transport protocol on satellite link in splitting TCP mechanism. Experiment results show that the proposal can significantly improve the TCP performance over satellite link.

Key words: TCP performance; splitting TCP mechanism; express transport protocol; error control mechanism; rate control mechanism

0 引言

TCP 协议是面向连接的传输层协议,向上层应用提供可靠的数据传输。随着通信事业的发展,卫星通信系统承载了大量的网络应用,其中基于 TCP 的应用大约占了应用总量的 85% 以上。由于 TCP 协议主要是以有线通信系统为应用环境设计的,没有考虑到卫星通信系统的特有属性,造成经由卫星链路传输 TCP 性能低下。目前提高卫星系统中 TCP 性能的方法主要有端对端方式和分段连接机制两种。端对端方式着重于 TCP 协议本身的改进,而分段连接机制则考虑将不同物理链路分割为不同的段,在每个段内采用不同的优化方式进行优化。在分段连接机制中,如何选择卫星链路部分的传输协议是整体性能优劣的关键。

XTP 是一种高速传输协议,一般应用于有线网络中,通过对其内部机制的深入研究发现它同样适用于卫星通信系统。

1 卫星通信系统对 TCP 性能的影响

卫星通信系统对 TCP 性能影响主要有以下三个方面原因:

1) 传播延迟

影响卫星网络延迟的主要因素是传播延时。以最为常用的同步卫星系统为例,其单向传播延迟为 250 ~ 280ms,双向传播延迟为 500 ~ 560ms^[1]。由于 TCP 采用滑动窗口机制,

TCP 的最大吞吐率决定于最大窗口尺寸和往返时间 (Round Trip Time, RTT) 时间的比值。在双向卫星系统中,标准 TCP (窗口的最大尺寸为 65 535) 的最大吞吐量为 $65\ 535 \times 8 / 0.55 = 953$ (Kbps); 其次,在 TCP 慢启动/拥塞控制阶段,均是以 RTT 为单位进行拥塞窗口调整,较大的 RTT 直接导致了在窗口的调整速度较低,限制了吞吐能力^[2-4]。

2) 信道错误率

卫星信道的比特错误率 (BER) 大约为 $10^{-3} \sim 10^{-9}$ 数量级^[1],这远远高于高速有线媒质 (如光纤),另外一些随机因素 (如雨衰、太阳黑子、微波干扰等) 使得信道出现突发错误。由于 TCP 无法区分报文的丢失是由网络拥塞造成的还是链路误码造成的^[5],在误码出现时, TCP 会进行不必要拥塞避免,降低了报文发送速率。另外, ACK 报文的丢失使吞吐量进一步恶化^[2-4]。

3) 信道不对称

考虑到大量 TCP 传输本身的较大非对称性 (如从 Web 服务器到远端主机)^[5],许多卫星系统在前向和反向信道间有较大的带宽不对称性,采用速度较慢的反向信道以节省卫星带宽。但严重非对称的配置会对 TCP 产生显著的影响。因为 TCP 的流量控制的基础是自同步机制^[2], ACK 的接收速率决定了数据报文的发送速率。反向信道的拥塞会造成对 TCP 在前向信道报文发送的抑制,对吞吐率产生负面影响^[4]。

收稿日期:2006-01-09

作者简介:吴结(1972-),男,北京人,硕士研究生,主要研究方向:计算机网络、嵌入式系统;高随祥(1962-),男,陕西人,教授,博士生导师,主要研究方向:图与网络流理论、算法设计与分析、网络最优化。

为了改善卫星 TCP 的性能问题,主要有两类解决方案:端对端方式和分段连接机制。端对端方式采用修改/优化主机 TCP 协议栈的方法改善性能,由于实施比较困难,因而分段连接机制是广泛应用的一种优化方案。

2 分段连接机制

分段连接机制改善卫星 TCP 性能的设计思路如下:如果端到端的链路由一些传输特性有很大不同的物理链路构成,那么将这些不同物理链路分割成不同段,在每个段内采用不同的优化方式进行优化,从而使得 TCP 的端到端总体性能得到较大的提高。分段连接机制是通过卫星协议网关(Satellite Protocol Gateway, SPG)实现的。

卫星协议网关是通过打断发送端与接收端的 TCP 连接,建立分段连接,在卫星链路中采用专门针对卫星链路设计的传输协议,以提高 TCP 在卫星链路中传输效率的网络设备。如图 1 所示,我们可以看到卫星协议网关位于卫星系统与地面系统连接处,来连接两类完全不同的物理链路。

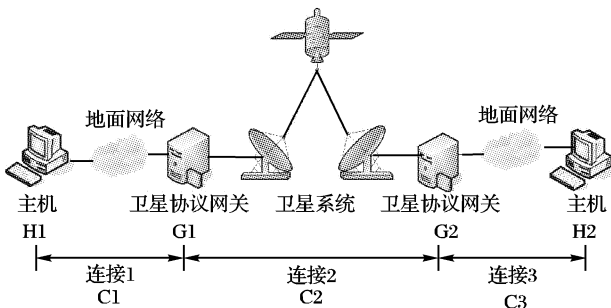


图 1 分段连接机制示意图

图 2 是卫星协议网关的协议栈示意图,从图中可以看到在卫星协议网关中专用协议替换了传统的 TCP 协议,如我们下面将要描述的针对 SCPC 体系的传输协议 SCPC-TP。通过这样的设计,首先,可以将卫星链路的缺陷隔离在有线路之外;其次,对应用完全透明,主机上的 TCP/IP 协议栈和应用不需要针对卫星协议网关进行任何额外的改动。

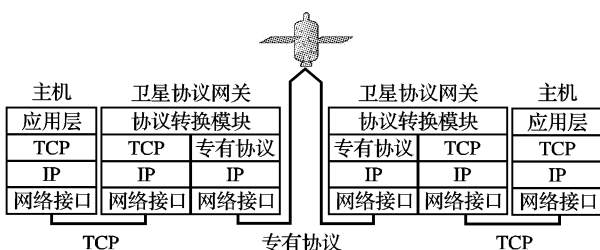


图 2 卫星协议网关协议栈示意图

分段 TCP 特点如下:

1) 分段 TCP 将端对端 TCP 连接分割成三部分,从而卫星链路中的问题不会与地面链路产生关联,因而,分段 TCP 方法支持采用局部解决方案来解决卫星链路中局部的问题,如:高延时、信道误码^[6]。

2) 分段 TCP 对应用完全透明,所有主机的 TCP/IP 协议栈和应用不需要为分段 TCP 进行任何修改和特殊设计。

3) 部署容易。卫星协议网关放置在卫星链路与地面链路之间,对原有的体系结构没有大的修改,实施非常方便^[7]。

为了获得理想的性能提升效果,在 C2 连接上采用何种传输协议是问题的焦点,选择的协议必须具备克服卫星通信系统缺陷的能力,同时必须保证数据的可靠传输。

3 快速传输协议(XTP)分析

XTP^[7]是一个通用的运输层协议,也是个轻型协议,它分离协议机制和控制策略,提供独立正交的协议机制,而由应用根据需要选择不同的控制策略。一般它被用于有线网络中,但是通过研究发现,快速传输协议同样也适用于卫星通信系统,可以克服卫星通信系统的一些缺陷。具体论述如下:

1) 64 位的序列号

XTP 与 TCP 一样采用滑动窗口机制,不同的是 XTP 使用 64 位的序列号来标识所有的数据,并使用 64 位的窗口尺寸。我们知道,单连接的吞吐量等于窗口尺寸与往返时间 RTT 的比值,相比 TCP 采用的 32 位序列号和 16 位窗口尺寸,XTP 在大时延高带宽的卫星通信系统中可以获得更高的吞吐量。

2) 隐式建立连接

XTP 支持连接的隐式建立。由发送端通过向接收端发送 FIRST 报文请求建立连接,当 FIRST 报文离开发送端时,发送端就直接进入“活跃”状态。发送端所需传输的数据既可以是后继的 DATA 报文,也可以直接包含在 FIRST 报文的信息段中。一旦这个 FIRST 报文到达接收端,如果 FIRST 报文携带的地址、通信量、选项信息符合接收端的要求,接收端也会进入“活跃”状态,表明双向的连接建立完成,可以允许上层应用进行双向的数据通信;如果不满足要求,接收端通过返回拒绝 DIAG 报文表示连接建立失败,相应的用户数据将由接收端直接抛弃。与 TCP 协议三次握手连接建立相比,XTP 可以通过一次报文交换完成建立连接、开启两个数据流的工作。

在 RTT 较大的通信系统中,可以通过采用隐式连接迅速建立连接,避免因建立连接而造成的带宽浪费。

3) 发送速率控制机制

在 TCP 机制中没有针对发送速率的控制机制,在滑动窗口允许的情况下全力发送数据。在 XTP 中引入了发送速率控制机制,发送端通过 RATE 和 BURST 两个参数和一个刷新定时器(RTIMER)对发送报文的速率进行控制,RATE 表示每秒可发送的最大字节数,BURST 则表示突发报文的最大的字节数,BURST 与 RATE 的比值就是 RTIMER 的超时周期。如图 3 所示,发送端保证在每个 RTIMER 超时周期内不发送大于 BURST 的数据,在每个 RTIMER 的定时间隔内都存在着“暂停”阶段,在此阶段将停止发送数据,这样就可以平稳地进行报文发送,不会造成很大突发报文。

针对带宽固定的卫星 SCPC 应用这是一项非常有用的机制,单路单载波(SCPC)方式是指每个载波只传送一路数据,两个相互通信站点之间占用卫星转发器的一对固定频率载波发射和接收,特别适用于大数据量的点到点实时通信,因而得到了非常广泛的应用。SCPC 方式最显著的特点是带宽固定,采用分段连接机制时,卫星链路部分采用具有速率控制机制的传输协议,由卫星协议网关将带宽分配给不同的连接,既可以避免链路中的拥塞,又可以实现接近于链路速率的吞吐量。

4) 选择重发的错误控制机制

在 XTP 中,错误控制机制被分为远端的 XTP 发送端和本地的 XTP 接收端两部分。如图 4 所示,在接收端,报文输入队列的状态通过三个状态参数表示:变量 dseq, req, hseq。它们都属于本地连接的上下文之中。这三个变量均是数据的序列号,只是表达意义不相同。变量 dseq 表示已经传递给用户的数据的下一个数据,所有序列号小于 dseq 的数据都已经被用户接收。变量 rseq 表示在已经被接收的数据中,与 dseq 连

续的无间隔的数据区间的最大序列号;变量 hseq 表示已经接收到的数据的最大的序列号,可能在 dseq 与 hseq 存在一些序列号间隙,这些间隙是由传输错误和拥塞丢弃报文造成的。这样如果接收到的数据连续没有错误并且都已经传递给用户,则 $dseq = rseq$;如果没有间隙则 $dseq = hseq$ 。

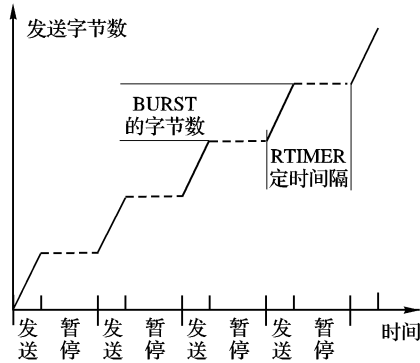


图 3 XTP 的速率控制机制

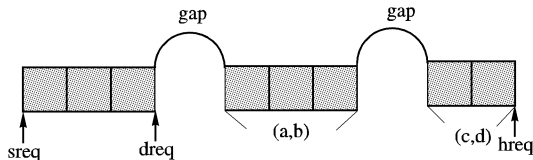


图 4 接收队列的状态表示

XTP 中的报文分为两类:DATA 报文和 CNTL 报文,DATA 报文负责承载数据,而 CNTL 报文为控制报文。在 CNTL 报文中记录了接收队列的状态,dseq 和 rseq 分别被放置在控制报文 CNTL 头部的 $\langle dseq \rangle$ 和 $\langle rseq \rangle$ 域中。但是 CNTL 报文不同于 TCP 中的 ACK 报文进行周期性地发送,它只在以下的两种情况下发送:

a) 当发送端有请求时,如果发送端需要了解接收端的状态,它可以在发送的报文中设置 SREQ 和 DREQ 标志,要求接收端通过 CNTL 报文报告其状态。

b) 倘若 XTP 的发送端选择了快速负确认机制,那么当接收端检测到接收队列中出现序列号间隔时,接收端会发送 CNTL 报文。但如果 XTP 的发送端没有设定快速负确认机制,接收端则不会发送 CNTL 报文。

在发送端接收到 CNTL 报文后会进行重传,与 TCP 不同的是,XTP 的重传是基于接收到的 CNTL 报文而不是依赖于类似于 TCP 定时器的超时。在 XTP 中主要的错误控制机制是回退 N 重传机制和选择性重传机制。

XTP 的回退 N 重传机制和选择重发机制共存于同一个基本的协议结构中。回退 N 重传机制与 TCP 中采用的重传机制相类似,从没有接收到第一个数据开始重发,不考虑已经收到正常数据。在 XTP 中使用了 CNTL 报文的 $\langle rseq \rangle$ 域,它用来指示需要进行重传的数据区间的开始序列号,发送端将对窗口内大于 $\langle rseq \rangle$ 域的数据进行重传。而选择性重传机制则利用了 CNTL 报文中的 $\langle nspan \rangle$ 和 $\langle spans \rangle$ 域,在 $\langle nspan \rangle$ 域中给出了序列号区间的数目, $\langle spans \rangle$ 域则是具体的序列号区间,这些序列号区间用来标记那些在 rseq 与 hseq 范围内,已经被正常接收的数据,如图 4 中所示 (a, b), (c, d) 区间,这样的区间可能会有多个,因而 $\langle spans \rangle$ 域的长度是可变的。发送端只对未确认的数据进行发送,避免了不必要的重传。

对于大时延和较高误码率的卫星链路而言,错误控制机制对性能产生很大的影响,在一般的高带宽、低时延、低误码的环境中,选择性重传机制与回退 N 重传机制具有大致相当

的性能。但在高误码、高时延的卫星通信系统的环境中,选择性重传机制的性能要明显优于回退 N 重传机制。

5) 同步握手机制

XTP 具有一种同步机制,命名为同步握手机制。每个连接上下文均有一个 sync 变量和一个 echo 变量。sync 变量是本地的计数器,在每发出一个 DATA 报文就会加一,echo 变量则是上一次与对端通信时对端的 sync 值。

在发送端,当前的 sync 被放置在每个发出报文的头部的 $\langle sync \rangle$ 域中,而 echo 则被放置在 CNTL 报文的 $\langle echo \rangle$ 域中。一次同步握手过程由 2 个 CNTL 报文的交换构成。首先,发送端发送 1 个 (CNTL, SREQ) 报文,该报文记录发送端当前的 $sync_1$ 和 $echo_1$ 的值;发送到接收端后,接收端会使用报文中的 $\langle sync_1 \rangle$ 域来更新当地的 $echo_2$ 变量。第二步,作为对 SREQ 请求的响应,当接收端生成 1 个 CNTL 报文回送到发送端时,CNTL 报文的 $\langle echo_2 \rangle$ 域设置为 $echo_2$ 变量,CNTL 报文到达发送端,进行报文的 $\langle echo_2 \rangle$ 域与本地的 $sync_1$ 变量的匹配,实现发送端与接收端的同步。在同步握手机制中,通过使用在 CNTL 报文中的 $\langle time \rangle$ 和 $\langle techo \rangle$ (time echo) 域实现 RTT 时间的测量。

由此可见:同步握手机制是一个控制报文的交换过程,通过这一机制使得发送端确切了解接收端的状态,可以消除发送端因为报文丢失和网络异常对接收端的状态做出不正确的判断,在出现大量误码情况下,同步握手机制的作用更为明显。

从表 1 可以看到,虽然 XTP 协议设计为高速网络传输协议,但相比 TCP 协议,它对卫星通信系统可以进行很好的处理,其完善的错误控制机制、隐式建立连接、发送速率控制和更高的序列号范围等机制都有利于克服卫星通信系统的缺陷,使之成为一种适合卫星通信系统的传输层协议。因而,我们建议将 XTP 协议作为分段连接机制中卫星链路中的传输协议。

表 1 XTP 与 TCP 的对比

比较内容	XTP	TCP
连接管理	隐式	三次握手
连接建立方式	二次握手	三次握手
错误控制	发送端要求	回复数据报文
重传策略	回退 N 重传,选择性重传	回退 N 重传
错误发现方法	快速负确认	RTO 超时
流量控制机制	滑动窗口	滑动窗口
速率控制机制	参数控制	无

4 测试

为了证明该建议的正确性,我们进行了性能测试。测试的目的是验证在分段连接机制中采用 XTP 协议后对端对端 TCP 性能的影响。我们的测试环境如图 5 所示。

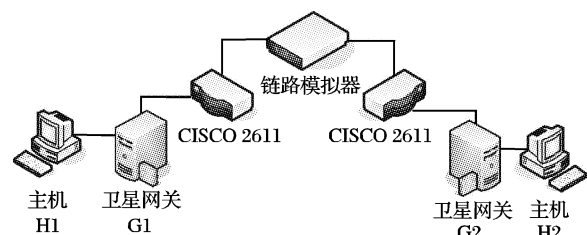


图 5 系统测试结构

主机 H1 和 H2 中采用的是 Windows 2000 Professional

(sp3) 和 Windows Server 2000 (sp3), 运行普通的 TCP/IP 协议栈(TCP Reno)。链路模拟器可以设置不同的速率、时延和误码率来模拟卫星链路。路由器为 Cisco 公司的 2611, 其端口最大的吞吐率为 8Mbps, 与链路模拟器的连接采用 v. 35 接口。卫星协议网关 G1 和 G2 实现分段连接机制的功能, 在与主机间的通信时采用 TCP 协议, 而两个网关之间也就是卫星链路则采用 XTP 协议。

为了考察单 TCP 连接的吞吐量, 我们采用 FTP 应用进行测试。H1 中安装 CuteFtp Pro, H2 中采用的 Ftp Server 为 Serv-U Ftp Server。分别关闭和打开卫星协议网关的分段连接功能, 采用不同速率和时延以及误码率进行测试。

1) 在 RTT = 560ms 时, 对标准 TCP 和采用 XTP 的分段连接(以下简称分段连接)按照不同的带宽进行 FTP 下载测试(下载文件大小为 10MB), 在图 6 中可以看到采用分段连接的带宽利用率可以达到 90%, 而采用标准 TCP 的带宽利用率, 由于窗口尺寸的限制, 随着带宽的增加越来越低。

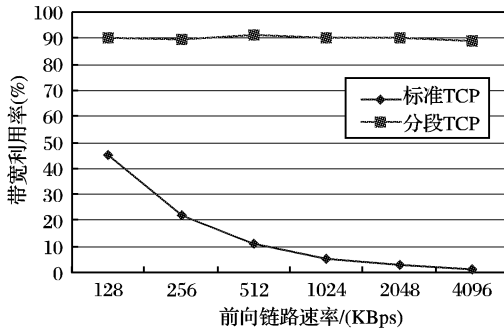


图 6 标准 TCP 和分段连接对比 (RTT = 560ms)

2) 在 RTT = 560ms 时, 采用分段连接按照不同的误码率进行 FTP 下载测试(下载文件大小为 10MB)。在图 7 中可以看到, 采用分段连接在误码率为 10^{-8} 的情况下能保证带宽的利用率 > 80%; 其次, 无论采用何种的误码率进行测试, 误码

率对高带宽影响要比低带宽影响大得多。

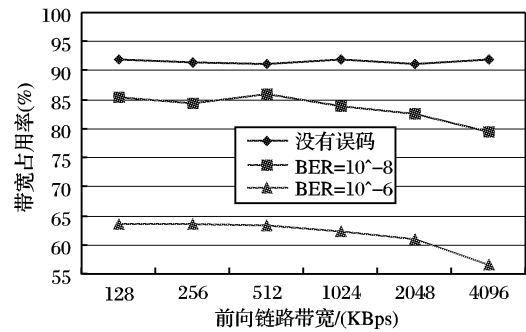


图 7 不同误码对分段连接的影响 (RTT = 560ms)

通过以上测试, 我们可以看到: 1) 采用建议的分段连接机制可以显著提高 TCP 端对端的性能, 使 TCP 链路的带宽利用率达到 90%。2) 在有误码的情况下仍有比较高的性能, 在误码率为 10^{-8} 的情况下仍能保证带宽的利用率大于 80%。

参考文献:

- [1] MARAL G. VSAT 网络[M]. 汲惠媛, 姜焕成, 译. 北京: 电子工业出版社, 1997.
- [2] JACOBSON V, KARELS MJ. Congestion Avoidance and Control [A]. ACM SIGCOMM[C], 1988.
- [3] ALLMAN M. Improving TCP Performance over Satellite Channels [D]. Ohio University, 1997.
- [4] BALAKRISHNAN H, SESHAN S, AMIR E, et al. improving TCP/IP Performance Over Wireless Networks[A]. ACM MobiCom[C], 1995.
- [5] STEVENS WR. TCP/IP 详解(第一卷)协议[M]. 任守奎, 译. 北京: 北京大学出版社, 1999.
- [6] IETF RFC 3135. Performance Enhancing Proxies Intended to Mitigate Link-Related Degradations[S], 2001.
- [7] TED S. XTP-98-20, Xpress Transport Protocol Specification Revision 4. 0[S]. XTP Forum, 1998.
- [8] 王秉均, 王少勇, 田宝玉, 等. 现代卫星通信系统[M]. 北京: 电子工业出版社, 2004.

(上接第 1553 页)

Step6.4.1.3 until 抽样稳定准则满足。

Step6.5 退温 $T = ST$, S 是退温速率。

Step6.6 更新“best so far”个体和适应度。

Step6.7 count = count + 1。

Step7 结束。

2 仿真结果

本文对 5 种实际网络进行了研究, 分别为美国国防部先进项目署光网络 (ARPANet)、英国光网络 (UKNet)、欧洲光网络 (EON)、中国高速信息示范网 (CAINONET) 和美国自然科学基金网 (NSFNet)。

表 1 均匀通信模式下各实际网络的最小波长需求 (备用路由数 4)

网络	理论边界	WP 波长需求		
		LP	SP	RP
ARPANet	33	33	33	33
UKNet	19	19	21	20
EON	18	18	20	19
NSFnet	13	13	14	13
CAINONET	12	12	13	13

表 1 给出了利用模拟退火遗传算法求得的在均匀通信模式(任意节点对间有且仅有一条光通道请求)下的各实际网络的最少波长需求数。算法的参数选择如下: 初始交配概率

0.7, 变异概率 0.1, 初始接受概率 0.1, 退火速率 0.95, 群体规模设为 30, 使用迭代次数 30 来控制遗传算法的结束。每个结果都是两次计算的最好解。从结果可以看出, 采用模拟退火遗传算法, 在采用长路优先的 FF 算法进行波长分配时, 优化结果达到最优, 和文献[4]相比, 有更好的优化效果。

仿真表明, 采用模拟退火遗传算法来解决静态 RWA 问题, 能取得较理想的优化解。

参考文献:

- [1] CHLAMTAC I, GANZ A, KARMI G. Lightpath Communications: An approach to high bandwidth optical WANs[J]. IEEE Transactions on Communications, 1992, 40(7), 1171 - 1182.
- [2] CHLAMTAC I, GANZ A, KARMI G. Lightnet: Lightpath. based solutions for wide bandwidth WANs[A]. INFOCOM'90[C], 1990. 1014 - 1021.
- [3] 李乐民. 光网络选路和波长分配研究[J]. 中兴通信技术, 2004, 11(6): 1 - 3.
- [4] 吴建华. 全光网络路由和波长分配技术研究[D]. 四川大学, 2003.
- [5] 叶亚斌. 利用遗传算法改进的路由和波长分配算法[J]. 清华大学学报(自然科学版), 2002, 42(7): 863 - 865.
- [6] BECKMANN D, KILLAT U. Routing and wavelength Assignment in Optical Networks Using Genetic Algorithm[J]. European Transactions on Telecommunications, 1999, 10(5): 537 - 544.