

卫星网络业务自相似特性分析

魏德宾^{1,2} 潘成胜¹ 韩睿^{1,2}

(辽宁省通信网络与信息处理重点实验室 大连 116622)¹ (大连大学信息工程学院 大连 116622)²

摘要 根据分包数据系统的星上数据源特点及随机过程理论,得到星上数据源的模型参数。进一步利用 STK 和 OPNET 软件建立卫星网络模型,依据数据源的建模结果为低轨卫星添加业务,并在高轨卫星处收集网络业务量。最后,分别运用 R/S 分析法、方差-时间图法和周期图法对收集到的卫星网络业务量的自相似程度和突发性的一个重要参数——Hurst 参数进行辨识。仿真结果表明卫星网络业务量也具有自相似性。

关键词 Hurst 参数,R/S 分析法,方差-时间图法,周期图法,卫星网络业务量

中图分类号 TN927 **文献标识码** A

Self-similarity Analysis of Satellite Network Traffic

WEI De-bin^{1,2} PAN Cheng-sheng¹ HAN Rui^{1,2}

(Key Laboratory of Communications Network and Information Processing of Liaoning Province, Dalian 116622, China)¹

(College of Information Engineering, Dalian University, Dalian 116622, China)²

Abstract The model parameters of data sources were obtained based on the characteristic of the packet onboard data system and stochastic process theory. Furthermore, the satellite network model was established by STK and OPNET. The model parameters of data sources were loaded to the LEO, and the network traffic was collected on the GEO. Lastly, R/S analysis, variance-time plots and periodogram-based analysis were applied to the Hurst parameter estimation of the degree of the self-similarity and burstiness of the satellite network traffic. Simulation results show that the satellite network traffic is characterized by self-similarity.

Keywords Hurst parameter, R/S analysis, Variance-time plots, Periodogram-based analysis, Satellite network traffic

在通信网络中,网络的流量特性是网络设计和性能分析中必须考虑的一个重要因素,在网络协议设计、性能优化和网络设备研究等方面具有至关重要的作用。

随着卫星技术的不断发展,卫星通信网络在通信领域中占据着越来越重要的位置。卫星网络是一个容量大、速率高的传输网络,它具有传输时延大、多普勒频移大、信道噪声大、拓扑结构高动态变化、空地通信频繁中断等特点,这些复杂的动态特性可以通过网络流量充分反映出来。因此,对卫星网络的流量进行研究是设计和分析卫星网络体系架构的切入点。根据系统的网络流量进行分析,找出其内在的特点,建立合理的网络流量模型,并根据模型进行流量预测,是对卫星资源进行控制和调配、改进和提高网络性能的一种重要手段。

长期以来,网络流量特性的研究受到研究人员的高度重视。目前对网络流量的研究主要集中在地面有线网络和无线局域网中,并发现这些网络业务在较大的时间尺度范围内具有自相似性^[1-6]。但由于卫星网络中的物理介质及传输协议的不同,地面网络的关于业务流量的分析结果不能直接应用于卫星网络。本文首先根据分包数据系统的星上数据源特点,运用随机过程的概念描述了星上数据源产生数据包的过

程,得到星上数据源的模型参数。其次,利用 STK 和 OPNET 仿真软件建立卫星网络拓扑结构。进一步将数据源的建模结果添加到低轨卫星节点处,并在卫星网络的骨干节点处收集网络业务量。最后,运用 R/S 分析法、方差-时间图法和周期图法对卫星网络业务自相似程度和突发性的一个重要参数——Hurst 参数进行辨识。仿真结果表明卫星网络业务量也具有自相似性。

1 星上数据业务的源包模型

对于分包体制的星载数据系统,源包是星上各应用过程与数据系统进行数据交换的单位,它的生成时间和数量均是可变的,根据星载数据系统的特点^[7],采用随机过程^[8]的概念来描述源包的产生过程。

设 $N(t)$ 表示某一应用过程在 $(0, t]$ 时刻随机产生的源包个数,则 $N(t)$ 是一个计数过程。此外,星上任一分系统可能包含若干个独立规律的子应用过程。可以认为,源包在任一极短时间区间 Δt 内,同一子应用过程不会同时产生两个源包,即在 $\Delta t \rightarrow 0$ 时,有

$$P\{N(t+\Delta t) - N(t) = 1\} = \lambda\Delta t + o(\Delta t)$$

到稿日期:2012-07-00 返修日期:2012-11-15 本文受国家 863 计划项目(2011AAXX),辽宁省教育厅重点实验室项目(LS2010007),辽宁省科技厅创新团队资助项目(LT2010007)资助。

魏德宾(1978—),男,硕士,讲师,CCF 会员,主要研究方向为卫星通信网络,E-mail:weidebin@163.com;潘成胜(1962—),男,博士,教授,CCF 会员,主要研究方向为导航制导与控制等;韩睿(1982—),女,硕士,讲师,CCF 会员,主要研究方向为卫星通信网络。

并且

$$\sum_{j=2}^{\infty} P\{N(t+\Delta t) - N(t) = j\} = o(\Delta t)$$

式中, $o(\Delta t)$ 是比 Δt 的高阶无穷小, λ 为单位时间内源包的平均到达率, $N(t)$ 是独立增量过程。

在众多的子应用过程数据源中, 按时间统计特性可分为两类: 平稳独立增量过程和时倚性独立增量过程。对于前者可以用强度为 λ 的齐次泊松过程来描述, 而后者采用非齐次泊松过程描述。

设源包依次出现的时刻 $t_1, t_2, \dots, t_n, \dots$ 是一强度为 λ 的泊松流, $\{N(t), t \geq 0\}$ 为相应的泊松过程。用下列记号 $w_0 = 0, w_n = t_n, n = 1, 2, \dots$ 表示第 n 个源包出现的等待时间, 则 w_n 是一随机变量, 其概率密度为

$$f_{w_n}(t) = \begin{cases} \frac{\lambda(\lambda t)^{n-1}}{(n-1)!} e^{-\lambda t}, & t > 0 \\ 0, & \text{其它} \end{cases}$$

即, 泊松过程的等待时间 w_n 服从 Γ 分布。又记 $T_i = w_i - w_{i-1}, i = 1, 2, \dots$ 。它被称为相继出现的第 $i-1$ 个源包和第 i 个源包的时间间隔, 则 $T_i (i = 1, 2, \dots)$ 的概率密度为

$$f_{T_i}(t) = \begin{cases} \lambda e^{-\lambda t}, & t > 0 \\ 0, & t \leq 0 \end{cases} \quad i = 1, 2, 3, \dots$$

即相邻两个源包产生的时间间隔是参数为 λ 的指数分布。

在对整个星上数据系统数据源建模的过程中, 把每一个用户产生的源数据序列进行分类, 分为若干个独立的子过程, 每个子过程按单一的规律产生源包。并且选择一种基本长度的源包作为基本源包, 把各种子过程的源包长度用相当于多少个基本源包的个数来表示。表 1 给出了卫星上部分信源的数据类型及其数据建模结果^[7]。

表 1 星上分包系统模型参数及取值

分系统	λ 取值和包个数 l 分布情况		
	健康数据	故障数据	遥控响应
姿轨控	$\lambda = 1.1$ $l = 20$	$\lambda = 10e^{-6}$ $l = 18000$	$\lambda = 0.1$ $l = 200$
热控	$\lambda = 1.03$ $l = 3$	$\lambda = 10e^{-6}$ $l = 1800$	$\lambda = 0.1$ $l = 80$
电源	$\lambda = 1$ $l = 20$	$\lambda = 10e^{-6}$ $l = 3000$	$\lambda = 0.1$ $l = 80$
总体电路	$\lambda = 1$ $l = 2$	$\lambda = 10e^{-6}$ $l = 1200$	$\lambda = 0.1$ $l = 80$
测控射频	$\lambda = 1.1$ $l = 5$	$\lambda = 10e^{-6}$ $l = 1200$	$\lambda = 0.1$ $l = 80$
数据管理	$\lambda = 1.1$ $l = 9$	$\lambda = 10e^{-6}$ $l = 2400$	$\lambda = 0.1$ $l = 100$
公共勤务	$\lambda = 1.1$ $l = 7$	$\lambda = 10e^{-6}$ $l = 1800$	$\lambda = 0.1$ $l = 80$

其中 l 表示基本包的个数, 一个基本包等于 10 字节。视频流业务模型参数及取值如表 2 所列^[9]。

表 2 视频流业务模型参数及取值

模型参数	服从分布	分布函数与参数
帧间隔时间	定值	100ms
包的大小	Pareto 分布	$F(x) = 1 - \left(\frac{\beta}{x}\right)^{\alpha}$ $\alpha = 1.2, \beta = 652$

设话音传输带宽业务为 64kbps, 采用泊松过程的 ON/OFF 模型。模型参数及取值如表 3 所列。

表 3 话音业务模型参数及取值

模型	分布	分布参数取值
呼叫持续时间	指数	120s
ON 状态时长	指数	0.352s
OFF 状态时长	指数	0.65s
包长	定值	160

图片业务采用等时间间隔 (0.83s), 包长为定长 8000 字节。

2 二阶自相似过程

2.1 二阶自相似过程的定义

设 $X = \{X_i; i = 1, 2, 3, \dots\}$ 为一个广义平稳离散随机过程, X_i 表示第 i 个时间间隔到达网络的数据包数。 X 具有恒定均值 $\mu = E\{X_i\}$ 和有限方差 $\sigma^2 = E\{(X_i - \mu)^2\}$, 其自相关函数为 $r(k) = E\{(X_i - \mu)(X_{i+k} - \mu)\} / \sigma^2, (k = 1, 2, \dots)$ 。随机过程 X 的长度为 m 的聚集过程 $X^{(m)}$ 定义为

$$X_i^{(m)} = \sum_{k=m(i-1)+1}^{mi} X_k / m, n = 1, 2, \dots$$

对每个 $m, X^{(m)}$ 都定义了一个广义平稳随机过程, 其方差和自相关函数分别为 $V^{(m)}$ 和 $r^{(m)}(k)$ 。

定义 1 如果随机过程 X 的自相关函数满足

$$r(k) = [(k+1)^{2-\beta} - 2k^{2-\beta} + (k-1)^{2-\beta}] / 2 \triangleq g(k), k \in Z_+$$

则称 X 为严格二阶自相似过程, 且具有 Hurst 参数, $H = 1 - \beta/2, 0 < \beta < 1$ 。

定义 2 如果随机过程 X 自相关函数满足

$$r(k) \sim ck^{-\beta}, k \rightarrow \infty$$

式中, c 为正常数; 则称 X 为长程相关过程, 且具有 Hurst 参数 $H = 1 - \beta/2, 0 < \beta < 1$ 。

定义 3 如果随机过程 X 的 m 阶聚集过程 $X^{(m)}$ 的自相关函数满足条件

$$\lim_{m \rightarrow \infty} r^{(m)}(k) = g(k), k \in Z_+$$

则称 X 为渐近二阶自相似过程, 且具有 Hurst 参数 $H = 1 - \beta/2, 0 < \beta < 1$ 。

2.2 自相似过程 Hurst 参数辨识

对于自相似过程, 仅用 Hurst 参数 H 就可描述其尺度伸缩特性, 因此, 对自相似性的检验主要就是对 Hurst 参数的估计。以下介绍求 Hurst 参数的 3 种方法。

2.2.1 R/S 分析法

对序列 $\{X_i, i \in Z^+\}$, 其部分和、样本均值和样本方差分别为

$$Y(n) = \sum_{i=1}^n X_i, \bar{X}(n) = \frac{1}{n} Y(n), S^2(n) = \frac{1}{n} \sum_{i=1}^n X_i^2 - \bar{X}^2(n)$$

定义 R/S 统计为

$$\frac{R(n)}{S(n)} \triangleq \frac{\max(0, W_1, \dots, W_n) - \min(0, W_1, \dots, W_n)}{S(n)}$$

式中, $W_k = Y(k) - k\bar{X}(n), (k = 1, \dots, n)$ 。若序列 $\{X_i\}$ 是自相似的, 则 $E\left\{\frac{R(n)}{S(n)}\right\} \sim cn^H, \forall n \rightarrow \infty$ 。其中, c 为正常数, H 为 Hurst 参数。因此, 图 $\ln E\{R(n)/S(n)\} \sim \ln n$ 应为一条直线, 其斜率为 H 。

2.2.2 方差-时间图法

聚集过程 $X^{(m)}$ 的方差 $V^{(m)}$ 与聚集阶数 m 之间满足如下条件: $V^{(m)} \sim cm^{-\beta}, m \rightarrow \infty$, 其中 c 为常数, $0 < \beta < 1$ 。对上式两边取对数得: $\ln V^{(m)} \sim -\beta \ln m + \ln c$ 。因此, 若序列 $\{X_i\}$ 为二阶

自相似,则图 $\ln V^{(m)} \sim \ln m$ 为一条直线,其斜率为 $-\beta$ 。通过关系式 $-\beta=2(H-1)$,就可以得到 Hurst 参数 H 的估计值。

2.2.3 周期图法

设序列 $\{X_i, i \in Z^+\}$ 是自相似过程 $\{X(t), t \in R\}$ 的离散序列。定义 $\{X_i\}$ 的周期图为

$$I_N(\omega) \triangleq \frac{1}{2\pi N} \left| \sum_{k=1}^N X_k e^{i\omega k} \right|^2$$

此函数可以作为过程 $\{X(t)\}$ 功率谱密度函数的一个渐近无偏估计,它的计算由下式给出

$$I(\omega_k) = \frac{1}{2\pi N} \left(\left| \sum_{i=1}^N x_i \cos(i\omega_k) \right|^2 + \left| \sum_{i=1}^N x_i \sin(i\omega_k) \right|^2 \right)$$

式中, $\omega_k = \frac{2\pi k}{N}$, $(k=1, 2, \dots, \frac{N}{2})$ 。将点 $(\omega_k, I(\omega_k))$ 绘在坐标系中,就得到了周期图 $I_N(\omega)$ 。若 $\{X(t)\}$ 为自相似过程,则 $\{X_i\}$ 的周期图应满足: $I_N(\omega) \sim |\omega|^{1-2H}$ 。因此,图 $\ln I_N(\omega) \sim \ln |\omega|$ 应该为一条直线,且斜率为 $\alpha=1-2H$ 的估计。

3 仿真及结果分析

在仿真过程中,设计了由 16 颗低轨卫星组成的 Walker Delta 星座和一颗东经 110 度赤道上空的地球同步卫星 GEO。星座的具体参数如表 4 所列。

表 4 星座参数

轨道参数	LEO
轨道高度(公里)	780
卫星数目	16
轨道数目	4
轨道倾角(度)	86.4°

根据这些参数在 STK 中设计星座,再将其导入 OPNET 仿真场景(见图 1)中,搭建网络仿真平台。仿真过程中,低轨卫星星座向地球同步卫星 GEO 发送数据。设每颗 LEO 卫星发送相同的业务。业务类型、包的发生率和大小见表 1—表 3,队列模型采用先进先出的 abc_fifo_ms 模型,发送速率为 120Mbps。

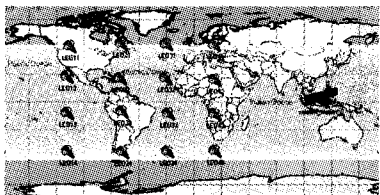


图 1 OPNET 仿真场景

设仿真时间为 6000s,选取整个卫星星座组成的网络作为仿真对象,并在地球同步卫星 GEO 处收集业务流量。每次收集的业务流量数据序列长度为 6000,序列中每个数据代表 1s 的时间尺度上通过网络的业务流量,图 2 为 GEO 处上行网络业务量。

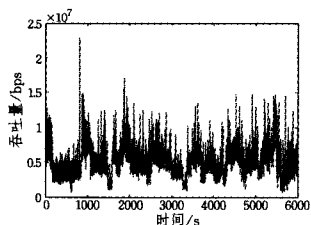


图 2 卫星网络 GEO 处上行网络业务量

下面对卫星网络流量的自相似性进行检验。

利用 R/S 分析法,画图 $\ln E\{R(n)/S(n)\} \sim \ln n$,并用最小二乘拟合为一条直线,其斜率为 Hurst 系数 $H=0.904$,见图 3。

利用方差-时间图法,画图 $\ln V^{(m)} \sim \ln m$,并用最小二乘拟合为一条直线,其斜率为 $-\beta=-0.169$ 。通过关系式 $-\beta=2(H-1)$,得到 Hurst 参数 H 的估计值为 0.915,见图 4。

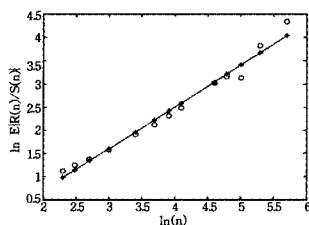


图 3 $\ln E\{R(n)/S(n)\} \sim \ln n$ 及最小二乘拟合

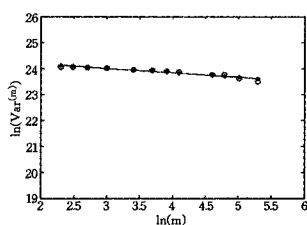


图 4 $\ln V^{(m)} \sim \ln m$ 及最小二乘拟合

利用周期图法,基于最小二乘法拟合对数坐标系下的周期图为一曲线,通过直线的斜率 $\alpha=-0.835$,可辨识出的 H 值为 0.917,见图 5。

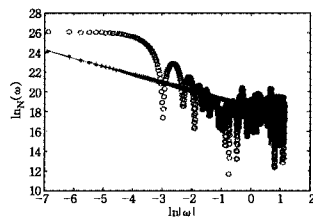


图 5 周期图及最小二乘拟合

用 3 种方法对卫星网络业务量的自相似参数估计结果如表 5 所列。

表 5 3 种 Hurst 参数估计方法及其估计结果

Hurst 参数估计方法	Hurst 参数估计结果
R/S 分析法	0.904
方差-时间图法	0.915
周期图法	0.917

结束语 自相似特性的发现使人们找到了实际网络流量与基于传统泊松和马尔可夫模型生成的网络流量的差距的原因,但这一发现是基于地面网络得到的。由于卫星网络拓扑结构的高动态特性,以及其物理介质和传输协议的不同,使得地面网络关于业务流量的分析结果不能直接应用于卫星网络。本文在对星载数据系统建模的基础上,利用 STK 和 OPNET 软件搭建卫星通信网络,在网络的骨干节点处收集网络业务量,并通过 3 种 Hurst 参数的辨识方法,得到了卫星网络业务量具有自相似特性的结论,这对卫星网络业务量的建模和预测将产生积极的意义。

参考文献

- [1] Leland W E, Taqqu M S, Willinger W, et al. On the self-similar nature of Ethernet traffic (extended version) [J]. IEEE/ACM Transactions on Networking, 1994, 2(1): 1-15
- [2] Beran J, Sherman R, Taqqu M S, et al. Long-range-dependence in Variable-bit-rate Video Traffic [J]. IEEE Transactions on Communications, 1995, 45: 1566-1579
- [3] Song S, Ng J K-Y, Tang B. Some results on the self-similarity property in communication networks [J]. IEEE Transactions on Communications, 2004, 52(10): 1636-1642

息的方式有两种,一种是定期采用 Get 命令来获取指定的能耗 MIB 信息;另一种是在设备的工作状态发生变化时(如端口速率由 1000Mbps 降为 100Mbps),使用 Trap 命令将相关信息发送给能耗管理系统。

②网络设备状态控制

网络能耗管理系统通过其他网络管理功能中获取的相关信息进行分析后,确定相关网络设备的工作状态,并通过 Set 命令对网络设备中的 EnergyStatus 信息进行控制,从而控制网络设备的工作状态。具体操作流程如下:

第 1 步 根据网络管理系统中其他功能的监控信息(如性能管理中的流量、负载信息)来确定相关网络设备的工作状态。如:路由器、交换机设备可根据端口的流量监测数据来调节端口的工作速率或关闭端口以节约电能消耗;

第 2 步 向网络设备发送 Set 命令,控制其状态控制 MIB 信息,使设备进入相关工作状态;

第 3 步 网络设备发送 Trap 命令,报告自己当前的工作状态信息;

第 4 步 能耗管理系统记录相关状态信息。

③网络能耗统计分析

网络能耗统计主要是集中在如何计算出单个设备的实际电能消耗和由能耗控制系统控制设备工作状态发生变化后节省的电能消耗。

定义集合 $S(S_1, S_2, \dots, S_n)$ 为设备的状态集合, P_i 为设备在状态 S_i 时的功率, T_i 为设备在状态 S_i 时的工作时长,则该设备消耗的电能 P_i 可用式(1)进行计算。

$$P_i = \sum_{i=1}^n P_i * T_i \quad (1)$$

此处的能耗特指网络设备运行期间消耗的能量,由于网络设备均使用电能作为能源,因此其能耗的单位为千瓦时。

考虑表 1 中设定的 MIB 信息时记录的是状态发生转变时的时间信息,能耗管理系统收集此信息后加入由能耗管理系统控制其转变的控制信息,即可形成工作状态转换时间表,如表 2 所列。

表 2 网络设备状态转换时间表

序号	设备工作状态	转变时间	是否由能耗管理系统控制转变
1	S_1	t_1	C_1
2	S_2	t_2	C_2
...
j	S_j	t_j	C_j
...
n	S_n	t_n	C_n

表 2 中, S_j 为状态集合 S 中的一种工作状态, t_j 为转换为当前工作状态时记录的时间, C_j 表示工作状态的转变是否由能耗管理系统来控制,取值为真或假。则设备在某一时间段上的电能消耗可用式(2)进行计算。

$$P_i = \sum_{j=1}^{n-1} P_{S_j} * (t_{j+1} - t_j) \quad (2)$$

(上接第 69 页)

[4] 于秦,毛玉明. WLAN 业务量特性分析及 Hurst 参数辨识[J]. 电子科技大学学报,2007,36(5):1074-1076,1113

[5] 赵佳宁,李忠诚. 基于模拟的网络流量自相似现象分析[J]. 计算机学报,2001,28(11):57-61

[6] 宋美娜,段云峰,宋俊德,等. 数据业务自相似特性研究[J]. 通信学报,2004,25(8):150-155

设备在某一时间段上由能耗管理系统控制所节省的电能消耗可用式(3)进行计算。

$$P_c = \sum_{j=1}^{n-1} (P_{S_j} - P_{S_{j-1}}) * (t_{j+1} - t_j) \quad (3)$$

其中的计算条件为 $P_{S_j} < P_{S_{j-1}}$ 且 C_j 为真。

由此,通过式(4)可以计算出网络中所有设备在某一时间段内的电能消耗 P_{TT} 。

$$P_{TT} = \sum_{i=1}^n P_i \quad (4)$$

通过式(5)可以计算出网络中所有设备在某一时间段内的所节省的电能消耗 P_{TC} 。

$$P_{TC} = \sum_{c=1}^n P_c \quad (5)$$

综上所述,在网络管理系统的能耗管理中通过以上能耗计算公式,可以对网络中的设备能耗和能耗节约进行定量的统计与分析。

结束语 通过对绿色计算和绿色网络的概念分析,结合网络管理的思想,提出了在网络管理中增加能耗管理的功能,并给出了具体实施方案,方案遵循简单易实施的原则,在网络设备中增加相关的能耗 MIB 信息,既保持了和原网络管理系统的完全兼容,又增加了绿色计算背景下对整个网络的能耗管理与控制,从而为全球的节能减排做出应有的贡献。

参 考 文 献

[1] 国务院新闻办公室.《中国应对气候变化的政策与行动(2011)》白皮书[EB/OL]. http://www.gov.cn/jrzq/2011-11/22/content_2000047.html,2011-12-08

[2] 林闯,田源,姚敏. 绿色网络和绿色评价:节能机制、模型和评价[J]. 计算机学报,2011(4):593-612

[3] Matthews E P, Nici J, Polonsky B, et al. Operating in the green: modeling eco-friendly telecom network management [J]. Bell LABS Technical Journal,2010,15(2):175-192

[4] 郭兵,沈艳,邵子立. 绿色计算的重定义与若干探讨[J]. 计算机学报,2009(12):2311-2319

[5] 过敏意. 绿色计算:内涵及趋势[J]. 计算机工程,2010(10):1-7

[6] 李麟. 服务器自适应节能系统介绍[J]. 信息技术与标准化,2008(9):16-19

[7] 吴甜,刘利祥,虎嵩林. 绿色数据中心的服务器节能机制与策略[J]. 微电子学与计算机,2011(8):108-111

[8] 唐广飞. 高性能路由器节能技术研究[D]. 长沙:国防科学技术大学,2006

[9] 葛雄资,冯丹,陆承涛,等. 绿色网络存储系统的动力学分析模型[J]. 计算机科学,2011(8):291-296

[10] Case J, Fedor M, Schoffstall M, et al. RFC1157: Simple network management protocol (snmp) [EB/OL]. <http://datatracker.ietf.org/doc/rfc1>,2011-12-17

[10] 鲍宁海,刘翔,张治中,等. WDM 节能光网络中的抗毁保护算法研究[J]. 重庆邮电大学学报:自然科学版,2012,24(3):278-282

[7] 顾莹琦,谭维织. 分包系统的星上数据源及其模型[J]. 遥测遥控,2001,22(2):26-31

[8] 王丽燕,沈玉波,刘洪. 随机过程[M]. 大连:大连理工大学出版社,2008

[9] Garrett M W, Willinger W. Analysis, Modeling and Generation of Self-Similar VBR Video Traffic [J]. ACM SIGCOMM Computer Communication Review,1994,24(4):269-280