

早期随机拥塞控制模型及其在 ATN 中的实现

张 军,高东岭,李忠孝,杜 明

(北京航空航天大学电子信息工程学院,北京 100083)

摘 要: 本文通过对传输控制协议/网际协议(TCP/IP)与面向连接传输层协议/无连接网络层协议(COTP/CLNP)中拥塞控制方法的比较分析,基于随机早期检测算法(RED)提出一种新的网络拥塞控制模型——早期随机拥塞控制模型(ERCCM),适用于所有采用绝对门限拥塞控制的网络系统。基于该模型和 RED 算法提出了一种适用于航空电信网(ATN)的随机早期置位方法(REM),并对该方法在 ATN 中的实现进行了仿真,仿真结果表明采用随机置位的方法可以解决绝对门限置位的三个主要问题,极大改善了网络的性能。

关键词: 航空电信网; 拥塞控制; 无连接网络层协议; 随机早期检测

中图分类号: TP393 **文献标识码:** A **文章编号:** 0372-2112 (2003) 03-0376-05

Early Random Congestion Control Model and the Realization in ATN

ZHANG Jun, GAO Dong-ling, LI Zhong-xiao, DU Ming

(School of Electronic Information Engineering, Beijing University of Aeronautics & Astronautics, Beijing 100083, China)

Abstract: A congestion control model called Early Random Congestion Control Model (ERCCM) is proposed based on the analysis of Random Early Detection (RED) and the comparison of basic congestion control method in Transfer Control Protocol/Internet Protocol (TCP/IP) and Connection Oriented Transfer Protocol/Connectionless Network Protocol (COTP/CLNP). It is suitable for all kinds of networks with absolute threshold congestion control. A method called Random Early Marking (REM) is also presented for use in Aeronautical Telecommunication Network (ATN) based on the principle of ERCCM and RED. The REM simulation results show that it can apparently overcome the defect of the absolute threshold marking methods in ATN and improve the performance of the network.

Key words: ATN; congestion control; CLNP; RED

1 引言

航空电信网(ATN)^[1]是国际民航组织(ICAO)提出的面向新一代航行系统应用的专用网络通信平台。航空电信网由空地子网和地-地子网两部分组成。空地子网提供空地间用户的通信服务,包括甚高频数据链(VDL)、卫星(AMSS)、高频数据链(HF)、二次监视雷达S模式以及舱门通信(GATELINK)等。地-地子网提供地面用户间的通信服务,可以采用分组交换(X.25)、帧中继(FR)、异步传输模式(ATM)以及各种IEEE/ISO兼容的802局域网协议等来实现。

在计算机网络中,当负载超过网络的转发能力时,就会出现拥塞。拥塞会造成数据包丢失、延时增大、网络资源利用率降低和网络吞吐量下降。航空电信网由于其特定的窄链路带宽的限制,使得拥塞控制显得尤为重要。

拥塞控制所采用的方法有很多, YANG 和 REDDY^[2]根据控制理论将所有的拥塞控制算法分为开环控制和闭环控制两类。开环控制是预防控制,它通过比较网络的承载能力和用户需要的网络资源来决定用户是否可以获得连接,在拥塞时

以丢弃数据包的方式来减轻网络压力。闭环控制基于反馈思想:(1)检测网络发生拥塞,比如门限超过,全程时间(数据发送到确认返回的时间间隔,又称RTT)过长,或产生互联网控制报文协议(ICMP)的源抑制报文;(2)通知相应设备有拥塞出现,采取纠正措施,如窗口机制;(3)相关设备调整(如源端减少发送量)消除拥塞,如重发定时器超时发送减半等。

随机早期检测算法(RED)^[3]是一种闭环控制方法,本文通过对该算法的分析研究,提出了一种早期随机拥塞控制模型,它适用于TCP/IP网络和无连接网络层协议(CLNP)^[4]网络等采用绝对门限进行拥塞控制的网络系统,并对其在ATN中的适用性做了研究和探讨。通过对绝对门限方法和随机早期置位(REM)方法的仿真比较分析,表明该模型可以极大地提高网络的性能。

2 ATN中的基本拥塞控制机制

2.1 网络层及传输层协议

ATN是基于OSI七层协议栈的端到端网络,其网络层采用ISO/IEC 8473无连接网络层协议(CLNP),典型的CLNP帧

结构如图 1 所示。

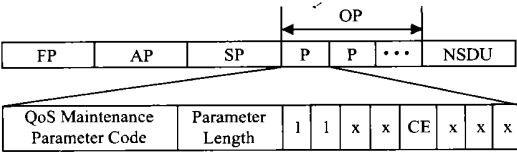


图 1 CLNP 帧结构

其中,FP 是 CLNP 固定封装格式部分,主要包括帧头长度、生命周期、网络层协议标识及版本号等内容;AP 是地址部分,包括源地址、目的地址及地址长度标识;NSDU 是数据部分,即为传输层的协议数据单元(TPDU);SP 是拆分部分,标明 TPDU 在原数据包中的偏置;OP 是可选部分,包括服务质量(QoS)维护参数码、参数长度标识及拥塞指示位 CE 等参数。

ATN 的传输层提供面向连接和无连接两种服务,分别采用 ISO/IEC 8073 面向连接传输层协议(COTP)^[5]和 ISO/IEC 8602 无连接传输层协议(CLTP)^[6]协议标准。表 1 给出了面向连接传输层协议的传输层协议数据单元(TPDU)的分类。

表 1 COTP 协议数据单元(TPDU)分类

简称	全称	中文定义
CR	Connection Request TPDU	连接请求
CC	Connection Confirm TPDU	连接确认
DR	Disconnect Request TPDU	断开请求
DC	Disconnect Confirm TPDU	断开确认
DT	Data TPDU	数据
ED	Expedited Data TPDU	加速数据(不需要连接建立)
AK	Data Acknowledge TPDU	数据确认
EA	Expedited Acknowledge TPDU	加速数据确认
RJ	Reject TPDU	拒绝连接
ER	Error TPDU	错误

2.2 明确流量控制机制

ATN 中,只有 COTP 能够提供拥塞控制服务。它采用明确流量控制机制(EFCE)实现拥塞控制,典型的数据传输流程如图 2 所示。

在数据传输之前,发送方发送连接请求(CR),如果接收方同意,则返回连接确认(CC),否则返回拒绝连接(RJ)。数据的传输过程是由发送方的数据(DT)和接收方的数据确认(AK)在双向传输构成的。完成传输后还要通过断开请求(DR)和断开确认(DC)的交互实现断开连接。连接确认(CC)和拒绝连接(RJ)中均有参数 CDT 指明连接时的物理缓冲大小,即最高窗口。数据(DT)的参数 PDU-NR、ROA 分别为序列号和要求 AK 确认的标识位。即明确流量控制既采用带序列号的确认过程,也实现窗口机制的接收端流量控制过程。图 3 给出了数据确认(AK-TPDU)的帧结构,参数 YR-TU-NR 和流量控制

确认(FCC)分别表示下一个希望接收的 TPDU 序列号,以及实际的窗口大小,即允许发送的 TPDU 的个数。

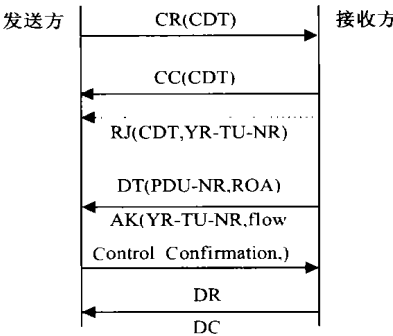


图 2 EFCE 的数据传输流程

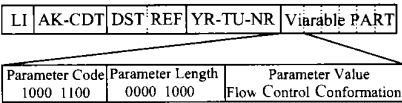


图 3 AK-TPDU 帧结构

2.3 网络层的 CE 标志

CLNP 帧结构即为 ATN 的网络层协议数据单元(NPDU), NPDU 头部的选项部分具有 QoS 维护参数,用以指定不同的服务质量。QoS 中有拥塞指示标识 CE,明确流量控制机制就是通过 CE 置位来实现的。CE 指示规则如下:

- (1) 在 NPDU 的发送方,该标志置为 0;
- (2) 中间路由器系统(IS)会对输出队列进行检测,当输出队列的深度大于 CE 置位队列长度时,超过部分的 NPDU 的 CE 被置为 1。

2.4 传输层广播窗口的更新

类似于 TCP/IP 的流控,ISO/IEC 8073 明确流量控制机制通过控制传输层的广播窗口进行流量控制。接收方的传输层实体通过 AK-TPDU 向发送方的传输层实体发送参数,用以指明其可接收的传输层协议数据单元的个数。发送方传输层实体根据此参数来调整广播窗口大小,称为广播窗口的更新。广播窗口的更新过程如下:

- (1) 收方向发送方给出预先设定的初始窗口 W_0 (首次为 W_0);
- (2) 到超过 W_0 的 TPDU 后,接收方进入采样阶段;
- (3) 采样阶段 N 和 NC 都置为 0 (其中, N 统计用来组成一个 NSDU 的 NPDU 的总数, NC 用来统计组成 NSDU 中 CE 标志被置 1 的 NPDU 的个数)。当 NC 超过 N 的 % 时,减少 W_0 (与小于 1 的参数相乘);反之,增加 W_0 ,增量为 ;
- (4) 产生 W_n ,由接收方给发送方;
- (5) 进入下一个更新过程。

3 RED 算法和 COTP/CLNP 网络与 TCP/IP 网络的拥塞控制。

3.1 TCP/IP 基本拥塞控制机制

TCP/IP 网络的拥塞控制机制^[7,8]需要三个参数:临界值、拥塞窗口和接收窗口。拥塞窗口指示实际发送的 TCP 分组的

个数;接收窗口指示接收端可以接收的物理缓冲的大小.慢启动过程从较小的拥塞窗口开始,没有超过临界值时以指数上升,超过临界值后以线性上升,重发定时器超时作拥塞窗口减半的处理.路由器检测到拥塞发生(如队列超过门限,RTT过大),发送源抑制包,源端收到后的处理与重发定时器超时相同.在此机制中,发生拥塞的指示出现在数据包在路由器的缓冲队列丢弃(超过物理缓冲长度),或 ICMP 源抑制报的产生.

3.2 RED 算法原理

以物理缓冲的长度作为丢弃门限的 TCP/IP 网络会发生许多问题^[9],如队列管理、全局同步和对突发业务的不公平.有许多方法可以改善 TCP/IP 网络的基本拥塞控制机制产生的问题,RED 是其中比较有效的一种.

RED 首次提出了采用线性方法实现早期随机控制,并广义地将拥塞指示称为标记(MARKING).对于确定的 TCP/IP 机制,标记强化为丢弃.

假定一个物理缓冲长度为 M 的路由器,RED 算法采用一个低通滤波器模型来计算平均队长的理论值 L_{avg} ,并将其与两个理论门限(T_{min} 和 T_{max})的比较,来确定丢包概率 P_a .每次丢弃数据包的概率大致与该连接在网关中所占用的带宽成比例. TCP/IP 中的 RED 算法描述如下:

```

For 到达数据包
    计算平均队长  $L_{avg}$ 
    if  $T_{min} < L_{avg} < T_{max}$ 
        计算丢包概率  $P_a$ 
        以概率  $P_a$  丢弃数据包
    else if  $L_{avg} > T_{max}$ 
        丢弃数据包
    
```

低通滤波器采用指数加权动态平均滤波,平均队长计算公式为:

$$L_{avg} = (1 - q) L_{avg} + q \cdot q \quad (1)$$

式中,权值 q 决定了低通滤波器的时间常数, q 为当前的队列实际长度.

标记概率 P_b 由平均队长线性决定:

$$P_b = P_{max} (L_{avg} - T_{min}) / (T_{max} - T_{min}) \quad (2)$$

其中, P_{max} 是当平均队长达到最大门限时丢包概率.

最终的丢包概率 P_a 随着计数器 (N_c) 增长:

$$P_a = P_b / (1 - N_c P_b) \quad (3)$$

其中, N_c 是上次丢弃后到达的包数.

图 4 给出了 RED 算法的平均队长与丢弃概率的关系.由平均队长与丢弃概率的关系可以看出,RED 算法主要有两个部分,一是计算平均队长,二是计算丢弃概率.由于采用低通滤波器模型,平均队长变化滞后于实际队长.它决定网关缓冲队列的理论长度,用于跟理论设定门限 T_{min} 、 T_{max} 进行比较和采取反馈控制.为了对所有数据包公平标记和避免全局同步(即网络吞吐量出现了震荡的情况),并且能控制平均队长,以随机的

图 4 RED 算法示意图

间隔来丢弃数据包. RED 算法采用低通滤波器来计算平均队长,因此由于瞬时拥塞而导致的队长短期增长并不严重影响平均队长.

3.3 COTP/CLNP 网络与 TCP/IP 网络的基本拥塞控制机制

TCP/IP 网络是以超过物理缓冲长度尾部丢弃引起重发定时器超时来指示拥塞的发生,在没有接收方的参与下由源端减少发送量. COTP/CLNP 网络发生拥塞的指示和 TCP/IP 网络中有所不同,它是出现在 NPDU 的头部的 CE 位的置位门限 T_m ,在接收方参与的窗口调节下,发送端减少发送数据量.

ATN 是 COTP/CLNP 网络,采用在中间路由器系统的接口进行门限置位的方法进行拥塞的检测与控制,该门限对应于 TCP/IP 网络的物理缓冲队列长度,是一种绝对化的门限.这种以绝对置位门限 T_m 进行控制的 ATN 会发生三种问题:(1)没有有效的队列管理方法.(2)网络吞吐量出现震荡,即全局同步.(3)突发业务由于其瞬时改变缓冲队列的长度明显,会出现对它的不公平,即该连接的业务占总吞吐量的比例不足.

4 拥塞控制模型分析

4.1 基本拥塞控制模型的引出

如图 5 所示的基本拥塞控制模型,其特点是绝对化的门限方法,拥塞的出现呈阶跃状的数字形式,可以指示拥塞发生的结果,不能指示拥塞发生的过程和预告.提出该模型,可以将控制过程数字化.对于 TCP/IP 网络采用门限的绝对化拥塞控制进行改善,RED 算法提出在绝对门限之前设立一段随机控制的缓冲,以解决绝对化的方法.

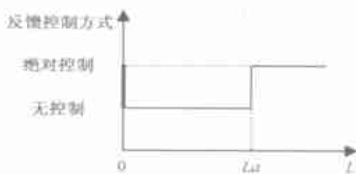
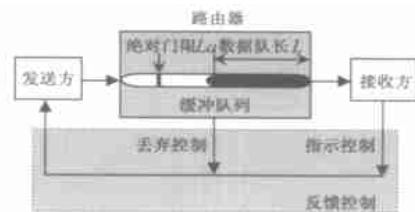


图 5 基本拥塞控制模型

4.2 早期随机控制模型的建立和分析

根据 RED 算法,本文提出了早期随机控制模型,在没有拥塞指示和绝对拥塞指示之间插入一段随机拥塞指示的缓冲.该模型适用于所有采用绝对门限方法的拥塞控制机制,并且可以在不同的协议基础上修改随机拥塞指示的方法,达到协议和网络性能的优化.其中 L_a 和 L_r 分别表示绝对门限和随机门限, L 为数据队长.如图 6 示,随机拥塞指示的设计,有效地解决了绝对化控制的三个主要的问题:(1)在绝对门限之前随机(按照各连接在总吞吐中的比例)进行拥塞指示,可以将最大流量的业务先减少,有效地控制缓冲队列的长度;(2)随机的拥塞指示使得部分连接的业务在出现拥塞预兆时减

少,而非到达绝对门限时全部减少,避免了全局同步;(3)突发业务(每次突发的数据量较大,链路延时和发送间隔较大的业务)容易出现暂时超过绝对门限的情况(但长时间平均值并不大),反馈控制使得该连接极大地减少窗口,即出现对它的不公平。如果随机拥塞指示的方法采用的是低通滤波器模型,如基于 RED 算法时,则由于其在计算队列长度时,突发业务不能明显地改变平均队长的大小,即不会出现拥塞指示和动作,因此可以有效的改善对突发业务的不公平。

对于早期随机控制模型中的随机控制部分,可以进行各种优化设计,比如抛物线置位,在需要网络适应突发性较弱的业务时会较线性更加有效。

4.3 随机早期置位(REM)算法

早期随机控制模型具有预防性检测和消除拥塞隐患的优点。模型在 ATN 中的实现—随机早期置位(REM)算法将 RED 的广义标记强化为置位并结合了 ATN 的明确流量控制机制(EFCE)。

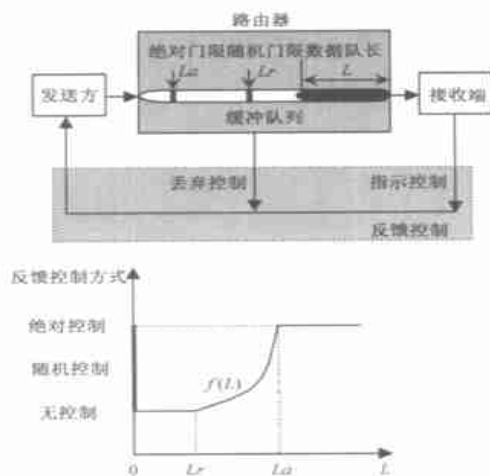


图6 早期随机控制模型

REM 算法采用置位方法对缓冲队列进行控制,在以物理缓冲长度为 M 的路由器中设置绝对置位门限 T_m 、最大随机置位门限 T_{max} 、最小随机置位门限 T_{min} ,通过平均队长 L_{avg} 与门限的比较来决定置位控制程度。最小随机置位门限之前无控制,最小和最大随机置位门限之间采取随机控制,绝对置缓冲长度之后则采取丢弃的控制方法。REM 算法描述如下:

```

For 批到达数据包
  计算平均队长  $L_{avg}$ 
  if  $T_{min} < L_{avg} < T_{max}$ 
    计算最终置位概率  $P_a$ 
    以概率  $P_a$  对数据包进行置位
  else if  $L_{avg} > T_{max}$ 
    对数据包进行置位
  
```

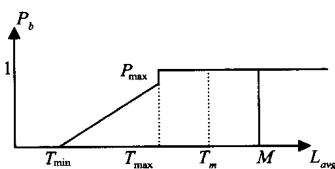


图7 REM算法示意图

else if $L_{avg} > M$

丢弃数据包

算法中平均队长的计算、置位概率 P_b 以及最终的置位概率 P_a 计算均和 RED 算法相同。平均队长 L_{avg} 与置位概率 P_b 的关系如图 7 所示。

5 仿真分析

REM 的一个改进是采用批数据包的概念。ATN 作为航空电信专用网络,与公用数据网的区别在于业务流量较少和网络资源配备较小。REM 算法中,对平均队长的计算的触发条件不采用 RED 中的每个到达的数据包,而是采用批到达的数据包。修改后的算法符合 ATN 的对资源要求较少的实际情况,但可能会引起计算过少无法及时控制的后果,需要对性能指标进行量化分析。

REM 算法对队列管理,全局同步和突发业务的不公平的改善等仿真参数列出如表 2 所示。

表2 仿真参数选取条件

假设条件	选择参数
(1) 路由器到各个节点的链路无拥塞	绝对门限置位 $M = 30k$, $\alpha = 0.5$, $\beta = 2$, NPDU 大小为 332 字节
(2) 处理时间和网络延时比忽略不计	随机早期置位 $T_{min} = 5k$, $T_{max} = 30k$, $T_m = 30k$, $w_q = 0.4$, $P_{max} = 0.6$, $M = 30k$, $\alpha = 0.5$, $\beta = 2$, NPDU 大小为 332 字节

5.1 对缓冲队列和延时的控制

为了简化分析,我们采用如图 8 所示的模型来代表网络。中间路由器系统 IS 连接 ES1、ES2、ES3 三个节点,ES1、ES2、ES3 到路由器的传输延迟分别为 $T_1 = T_3 = 5ms$, $T_2 = 15ms$ 。对 REM 算法进行队列控制的仿真,其结果如图 9 所示。图中,REM 表示随机置位的方法;TH 表示基本的绝对门限的方法。

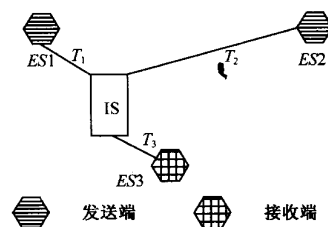


图8 简化网络模型

可以看出,缓冲队列的长度在物理缓存为 64Kbytes 时,基本的绝对门限方法几乎达到极限,而随机检测的早期置位方法有效地控制了流量,使得平均缓冲队列很小,延时相应地大大减小。

5.2 全局同步问题的解决

衡量全局同步问题的仿真参数仍取 5.1 节参数值,主要性能指标选为在间隔时间内的队列长度变化次数。如图 10,

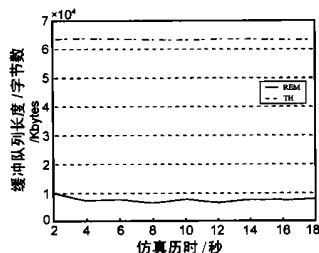


图9 缓冲队列和延时控制的仿真曲线

可以看出传统的方法每秒的抖动次数在 30 次左右,使用 REM 后减小了一半,基本控制在 20 次以下,有效地解决了全局同步问题。

5.3 对突发业务不公平的改善

如图 11 的模型,中间路由器系统 IS 连接 ES1、ES2、ES3、ES4 四个源节点和一个目的节点 ES5,各个源节点的发送窗口根据广播窗口随机调整。仿真采用的评价指标为 ES1 发送方的业务占整个队列的百分比,它衡量 ES1 对中间路由器系统 IS 的缓冲队列的贡献。选取传输延迟 $T_1 = 15\text{ms}$, $T_2 = T_3 = T_4 = T_5 = 5\text{ms}$ 。仿真结果如图 12 所示,图中看出随机置位方法使得突发业务源端 1 的贡献上升了约 1 倍,突发业务在总业务的比例从 1.5% 上升到约 3%。这表明 REM 减少了瞬时拥塞的影响,改善了对突发业务的不公平。

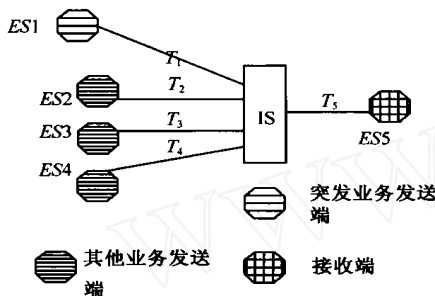


图 11 突发业务源模型

6 结论

本文在分析 RED 算法的基础上,提出了适用于采用绝对门限拥塞控制方法的系统的通用模型—早期随机拥塞控制模型。该模型

可以应用于 TCP/IP 网络,也可以应用于 COTP/CLNP、SPX/IPX 等其他网络。

针对该模型在 ATN 的应用需求提出的 REM 算法,可解决航空电信网没有有效拥塞控制机制的问题。仿真结果表明,REM 使得 ATN 中面临的三个主要问题的技术性能有明显的改善。

基于 CLNP 协议的 ATN,有区别于 TCP/IP 网络的许多特

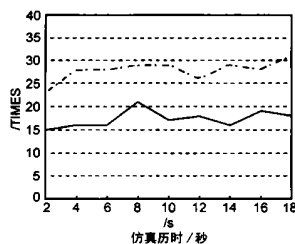


图 10 对全局同步问题的仿真曲线

点,如在选项部分有明确的 QoS 参数、安全参数和层优先级问题。对于文中提出的 REM 算法,未来待解决的主要问题包括:各门限的优化,明确流量控制机制和 REM 算法的参数配合等。模型中的早期随机拥塞指示方法,也不局限于线性方法,可以借鉴其他网络优化算法进行设计和试验。

参考文献:

- [1] Manual of Technical Provisions for the Aeronautical Telecommunication Network (ATN) (3rd Edition) [S]. 1997.
- [2] Cui-Qing Yang, et al. A taxonomy for congestion control algorithms in packet switching networks [J]. IEEE Network, 1995, 9(4): 34 - 45.
- [3] Sally Floyd, Van Jacobson. Random early detection gateways for congestion avoidance [J]. IEEE/ACM Transaction on Networking, 1993, 1(4): 397 - 413.
- [4] ISO/IEC 8473 - 1, Information Technology-Protocol for Providing the Connectionless-Mode Network Service: Protocol Specification [S].
- [5] ISO/IEC 8073, Information Technology-Open Systems Interconnection-Protocol for Providing the Connection-Mode Transport Service, ISO/IEC [S].
- [6] ISO/IEC 8602, Information Technology-Protocol for Providing the OSI Connectionless-Mode Transport Service [S].
- [7] D D Clark. Window and Acknowledgement Strategy in TCP [R]. RFC, IEPF 813, 1982.
- [8] J Nagle. Congestion control in IP/TCP Internetworks [R]. RFC, IEPF 896, Jan, 1984.
- [9] V Paxson, et al. Known TCP Implementation Problems [R]. RFC, IEPF 252, 1999.

作者简介:



张军男, 1965 年出生于安徽省合肥市, 1990、2001 年分别获北京航空航天大学工学硕士、博士学位, 现为北京航空航天大学教授, 博士生导师, 电子信息工程学院副院长, 主要研究方向: 航空数据通信, 空地一体化信息传输, 数据链技术, 新航行系统, 遥控遥测等。



高东领男, 1974 年出生于广东省, 1997 年毕业于广东工业大学, 毕业后从事电信技术工作, 熟悉数据通信网和路由交换技术。1999 - 2002 年在北京航空航天大学电子工程系攻读硕士学位, 研究方向为航空电信网 (ATN) 和数据链技术。