

ATM 网络中优先级业务漏桶算法的研究¹

蒋志刚 李乐民

(电子科技大学光纤通信国家重点实验室 成都 610054)

摘 要 本文对优先级业务的漏桶算法进行了分析, 得出了各种优先级业务的漏桶性能与输入业务统计特性和漏桶参数之间的关系。文中引入低优先级业务的固定服务速率调节因子 η , 并研究了它对高优先级业务和低优先级业务性能的影响。

关键词 ATM 网络, 优先级业务, 漏桶算法

中图分类号 TN913.2

1 引 言

ATM 网络具有灵活性强, 频带利用率高等特点, 能够支持不同种类的业务 (如数据、语音、图象等) 和不同服务质量要求的业务 (包括各种优先级业务)。由于 ATM 网络高速的信息传输和交换, 使得传统分组交换网络的流量控制方法不再适合于 ATM 网络。通常 ATM 网络采用预防式控制。漏桶法是进行流量控制的一种有效的方法。文献 [1-3] 对多种业务的漏桶算法进行了研究, 但对优先级业务的漏桶算法讨论的很少。优先级控制是 ATM 网络拥塞控制重要的功能组成部分。通常虚漏桶法是对漏桶溢出的分组进行标签, 当沿连接虚通道的各节点没有发生拥塞时, 这些标签的分组顺利通过, 当其中任一节点处于拥塞状态时, 这些标签的分组首先被丢掉。为了对不同优先级的分组合理地标签, 有必要对优先级业务的漏桶算法进行研究。

2 业务模型

图 1 是漏桶输入业务的模型。N 路独立同分布的信源经统计复用的分组流进入漏桶排队缓冲器。每路信源的流体流模型由 k 种优先级 (S_1, \dots, S_k , S_1 对应最高优先级, S_k 对应最低优先级) 的通-断信源组成。当通-断信源处于通 (ON) 状态时, 分组以一定速率周期地产生; 当处于断 (OFF) 状态时, 无分组产生。通-断信源是一种突发信源, 到达分组存在一定的相关性。为了分析的简明性起见, 我们假设各种优先级业务的转移概率强度矩阵相同, 由下式给出

¹ 1994-01-04 收到, 1995-08-22 定稿
国家自然科学基金资助项目

$$A = \begin{bmatrix} -\alpha & \alpha \\ \beta & \beta \end{bmatrix}, \quad (1)$$

则信源利用率为

$$p = \alpha / (\alpha + \beta). \quad (2)$$

处于 ON 状态的信源路数 i 的稳态分布为

$$\pi_i = \binom{N}{i} p^i (1-p)^{N-i}, \quad i = 0, 1, \dots, N. \quad (3)$$

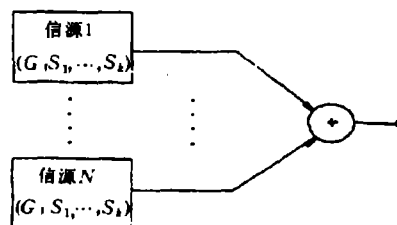


图 1 漏桶输入业务模型

3 前 k 种优先级业务总的排队分布及其漏桶性能

由于高优先级业务的排队性能不受低优先级业务排队的影响, 我们首先对前 k 种优先级业务总的排队分布进行分析。

设 $F_i(q(k))$ 为前 k 种优先级业务的虚队列排队长度 Q_f 和输入业务所处状态 I 的联合概率分布函数; $\gamma_i(k)$ 是有 i 路信源处于 ON 状态时, 前 k 种优先级业务的排队分组的变化速率为

$$\gamma_i(k) = v(i, k) - R, \quad (4)$$

其中 $v(i, k)$ 为有 i 路信源处于 ON 状态时, 前 k 种优先级业务的分组到达速率。

采用流体流法对基于虚排队漏桶模型的优先级业务的漏桶算法进行分析^[4,5], 可求出前 k 种优先级业务总的排队分布和实际排队队列占满的概率 P_{fi} , 则前 k 种优先级业务的分组丢失率为

$$P_{\text{loss}}(k) = \left(\sum_{i=1}^N \gamma_i(k) P_{fi} \right) / [Npv(k)], \quad (5)$$

其中 $v(k)$ 为每路信源处于 ON 状态时前 k 种优先级业务的分组到达速率。

前 k 种优先级业务的平均延时为

$$\bar{W}(k) = \bar{Q}_r(k) / [(1 - P_{\text{loss}}(k))Npv(k)], \quad (6)$$

其中 $\bar{Q}_r(k)$ 为前 k 种优先级业务实际排队的平均排队队长。

4 第 $k+1$ 种优先级业务的排队分布及其漏桶性能

(1) **优先级业务的服务规则** 无高优先级分组和低优先级分组时, 无分组被服务; 无高优先级分组有低优先级分组时, 对低优先级分组服务; 有高优先级分组时无论有无低优先级分组, 总是对高优先级分组服务; 对于同一优先级的分组采用先到先服务的服务规则 (FCFS)。

(2) 优先级业务的排队规则 当漏桶排队缓存器占满时, 此时到达分组的优先级如果是排队缓存器中分组优先级最低的, 该分组被丢失, 否则排队缓存器中优先级最低的分组被丢失, 依次类推。

可见第 $(k+1)$ 种优先级业务的排队分布只与前 k 种优先级业务的状态有关, 而与之优先级较低的业务的状态无关。设 $q(k)$ 表示前 k 种优先级业务排队的状态, q_{k+1} 表示第 $k+1$ 种优先级业务排队的状态, 第 $k+1$ 种优先级业务的排队分组的变化速率 $\gamma_{i,k+1}$ 为

$$\gamma_{i,k+1} = \begin{cases} iv_{k+1} - \max[0, R - v(i, k)], & q(k) = 0; \\ iv_{k+1}, & q(k) \neq 0. \end{cases} \quad (7)$$

考虑联合概率分布函数 $F_i(q_{k+1}, q(k))$

$$\begin{aligned} F_i(q_{k+1}, q(k)) &= P_r\{Q_f \leq q_{k+1}, Q_k = q(k), I = i\}, \\ 0 \leq i \leq N, \quad 0 \leq q(k) \leq B + M, \quad 0 \leq q_{k+1} \leq B + M - q(k), \end{aligned} \quad (8)$$

根据全概率公式有

$$F_i(q_{k+1}, q(k)) = \sum_{q(k)=0}^{B+M} P_r\{Q_k = q(k)\} F_i(q_{k+1}|q(k)). \quad (9)$$

对于前 k 种优先级业务任意的排队状态 $q(k)$, 可以推导出第 $k+1$ 种优先级业务的排队方程, 它与前 k 种优先级业务的排队方程相似, 只是排队方程中排队分组的变化速率 $\gamma_i(k)$ 变为 $\gamma_{i,k+1}$, $F_i(q(k))$ 变为 $F_i(q_{k+1}|q(k))$; 相应的边界条件为

$$\begin{cases} F_i(B + M - q(k)|q(k)) = \pi_i, & i \in \Omega'_-; \\ F_i(0|q(k)) = 0, & i \in \Omega'_+; \end{cases} \quad (10)$$

其中 Ω'_+, Ω'_- 是第 $k+1$ 种优先级业务所对应的状态区域:

$$\Omega'_+ = \{i | \gamma_{i,k+1} > 0\}, \quad (11)$$

$$\Omega'_- = \{i | \gamma_{i,k+1} \leq 0\}. \quad (12)$$

第 $k+1$ 种优先级业务排队方程的求解过程与前 k 种优先级排队方程的求解过程相似, 采用数值计算方法求取对应矩阵的特征根和特征向量, 通过解满足边界条件的线性方法组求取系数向量, 从而获得了分布 $F_i(q_{k+1}|q(k))$ 的解, 第 $k+1$ 种优先级业务排队队长的分布为

$$F(q_{k+1}) = \sum_{i=1}^N \sum_{q(k)=1}^{B+M} F_i(q_{k+1}, q(k)). \quad (13)$$

最后根据类似 (5), (6) 式的求法可求得第 $k+1$ 种优先级业务的分组丢失率 P_{loss}^{k+1} 和平均延时 \bar{W}_{k+1} 。

当高优先级业务的排队队长不等于 0 时, 低优先级业务的服务速率为 0, 这样低优先级业务的性能有可能变得很差而不能满足低优先级业务的性能要求。为了改善低优先级业

务的性能, 可以给低优先级业务分配一定量的固定的服务速率 R_L , 下面我们以双优先级业务为例来讨论这个问题。

我们定义低优先级业务的固定服务速率调节因子 η 为

$$\eta = R_L / (Npv_L), \quad \eta \geq 0, \quad (14)$$

其中 v_L 是低优先级业务处于 ON 状态时的分组到达速率, 则对应于 (4), (7) 式的高优先级业务和低优先级业务的排队分组的变化速率分别为

$$\gamma_{i,H} = i \times v_H - (R - \eta Npv_L), \quad (15)$$

$$\gamma_{i,L} = \begin{cases} i \times v_L - \max[\eta Npv_L, \eta Npv_L + R - \eta Npv_L - i \times v_H], & q(k) = 0; \\ i \times v_L - \eta Npv_L, & q(k) \neq 0. \end{cases} \quad (16)$$

高优先级业务和低优先级业务的排队方程及其求解过程与前类似, 这里就不再赘述了。我们在数值计算部分可以看到, 选择适当的调节因子 η , 可以在对高优先级业务性能影响不大的条件下, 使低优先级业务的性能大为改善。

5 数值计算和结论

双优先级业务是综合业务网络中较为常见的情况; 我们在这部分给出双优先级业务的漏桶性能的数值计算和模拟结果, 并就调节因子对高优先级业务和低优先级业务性能的影响进行讨论。

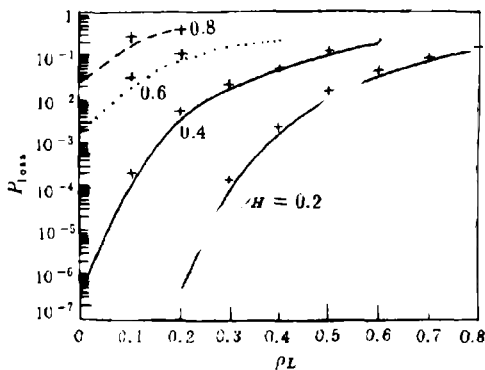


图 2 不同高优先级业务负荷情况下低优先级业务的分组丢失率随低优先级业务负荷变化的曲线

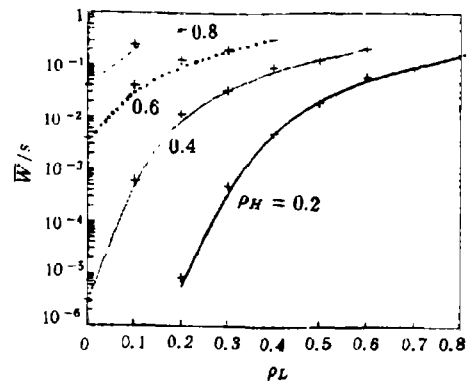


图 3 不同高优先级业务负荷情况下低优先级业务的平均延时随低优先级业务负荷变化的曲线

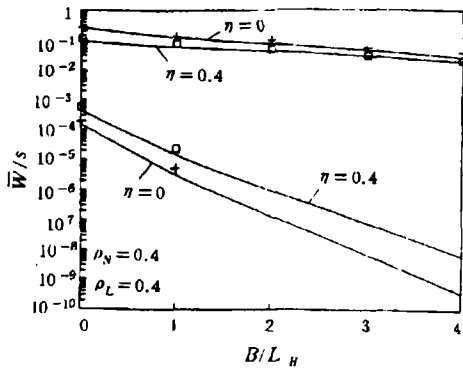


图 4 不同调节因子情况下高优先级业务和低优先级业务平均延时随参数 B 的变化

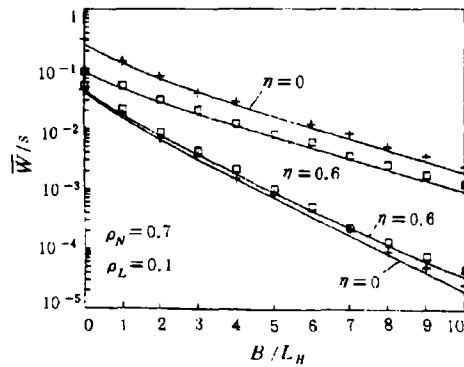


图 5 不同调节因子情况下高优先级业务和低优先级业务平均延时随参数 B 的变化

设高优先级业务和低优先级业务的统计参数为 $\alpha = 0.4, \beta = 1.0$ ；输入信源数 $N = 10$ ；高优先级业务处于 ON 状态时的归一化分组到达速率 $v_H = 1.0$ ，则其平均突发长度为 $L_H = (1/\beta)v_H = 1.0$ 。图 2，图 3 分别是在不同高优先级业务负荷情况下，低优先级业务的分组丢失率和平均延时随低优先级业务负荷变化的曲线。从图中可以看到在排队缓存器容量有限的情况下，低优先级业务的漏桶性能主要受总的业务负荷的影响。高优先级业务对低优先级业务的性能有一定的影响，高优先级业务的负荷相对低优先级业务的负荷越大，低优先级业务的性能越差。图中“+”，为计算机模拟值。

图 4，图 5 分别是在不同负荷及调节因子的情况下高优先级业务和低优先级业务的平均延时随漏桶令牌存储器容量 B (以平均突发长度 L_H 为单位) 变化的曲线。图中“+”，“□”号为计算机模拟值，它与理论计算值是吻合的。从图中可以看出，当低优先级业务的负荷相对高优先级业务的负荷较小时，采用较大的调节因子，可以在对高优先级业务性能影响不大的情况下，使低优先级业务的性能显著地改善，这正好可以克服高优先级业务对低优先级业务造成的不利影响。

参 考 文 献

- [1] Chuch M C, Cruz R L. Approximate analysis of average performance of (σ, ρ) regulators, IEEE INFOCOM'90, 1990, 874-880.
- [2] Berger A W. Performance analysis of a rate control throttle where tokens and jobs queue, IEEE INFOCOM'90, 1990, 30-38.
- [3] Yin N, Hluchyj M G. Analysis of the leaky bucket algorithm for on-off data sources, IEEE GLOBE-COM'91, 1991, 9. 2. 1-9. 2. 7.
- [4] Tucker R C F. IEEE Trans. on COM, 1988, COM-36(4): 479-483.
- [5] Anick D, Mitra D, Sondhi M M. Bell System Tech. J, 1982, 61(8): 1871-1894.

ANALYSIS OF THE LEAKY BUCKET ALGORITHM FOR PRIORITY SOURCES IN ATM NETWORKS

Jiang Zhigang Li Lemin

(University of Electronic Science and Technology of China, Chengdu 610054)

Abstract The leaky bucket algorithm for priority sources in ATM network is derived and the analytical expressions of the relation among the leaky bucket performance, the leaky bucket parameters and the statistical parameters of input traffic are obtained. The adjustment factor of constant serving rate for lower priority service, η , is introduced and its effect on the performance of the higher priority and the lower priority services are studied.

Key words ATM network, Priority sources, Leaky bucket algorithm

蒋志刚：男，1964年生，助理研究员，博士，从事宽带综合业务数字网的网络技术和拥塞控制的研究。

李乐民：男，1932年生，教授，博士生导师，从事信息传输和通信网方面的研究。