

一种基于对策模型的 ATM 网络连接接纳控制策略¹

陈惠民 王 普 李衍达

(清华大学自动化系 北京 100084)

摘 要 本文先简述了 ATM 网络进行连接接纳控制 (Connection Admission Control: CAC) 的主要方法, 其中主要综述了基于动态带宽分配的 CAC 策略, 而后从合作对策模型的角度讨论业务间共享链路资源的公平性问题, 提出一种基于时延带宽积的业务收益函数形式, 并通过遗传算法求解待优化的对策函数, 以决定对呼叫请求的接入或拒绝, 仿真结果表明此方法能够更好地保证不同带宽和服务质量要求的业务共享网络资源的公平性。

关键词 连接接纳控制, 合作对策模型, 遗传算法, 动态带宽分配, ATM 网络

中图分类号 TN913.24

1 引 言

异步传输模式 (Asynchronous Transfer Mode: ATM) 已逐渐成为宽带综合业务数字网 (B-ISDN) 的传输标准。ATM 网络以信元为单位传输, 允许不同业务流的信元在传输过程中通过统计复用提高网络链路的带宽利用率, 而且能够保证各类业务不同的服务质量 (Quality of Service: QoS) 要求, 并对网络可能出现的拥塞实施流量控制。ATM 网络拥塞控制分为预防式控制和反应式控制: 预防式控制包括连接接纳控制 (Connection Admission Control: CAC) 和带宽管制 (Bandwidth Enforcement)。前者在用户呼叫请求时决定是否为其所要传输的业务建立连接和分配网络带宽资源, 这主要依据用户提供的业务量特性、QoS 要求和网络当前的资源利用情况来限制接入业务数量; 后者用来限制被接入的业务按照它在呼叫请求之际得到的带宽来传送数据。反应式控制则通过发送反馈信元来调整业务端的发送速率, 防止拥塞。以下我们仅讨论用于 CAC 的带宽分配准则。

早期 ATM 网络带宽分配策略主要采用静态方法, 即在呼叫建立阶段, 用户向网络提供两组参数, 一组是指明其 QoS 要求 (如信元丢失率和最大传输时延)、另一组是其预期的业务量统计特性。如果网络具有足够的剩余带宽和缓冲资源, 在不破坏现有业务连接 QoS 的前提下, 网络将接入新业务并为其分配所需资源, 否则它将拒绝该呼叫请求。但是, 上述静态方法由于依赖于用户提供的业务量参数而存在以下问题^[1]: (1) 目前, 如何选择有效的业务源模型参数以便进行带宽分配还未达成共识; (2) 通常, 用户在呼叫阶段不可能精确估计出其业务量特性; (3) 一旦接入业务与其它信元流复用, 它的统计特性将发生改变, 因此由用户提供的相同业务参数不可能对此连接经过的所有交换节点均具有同样的精确性。

针对这些问题, 解决的途径之一是要求用户仅提供一组简单的参数, 如峰值位率和平均位率, 然后网络通过假定“最坏”业务模式来进行保守的资源分配, 各种业务模式估计方法的综述见文献 [2], 尽管此法简单, 但单独的保守机制将导致网络资源的利用率低。因此有

¹ 1997-02-24 收到, 1997-12-03 定稿
国家自然科学基金资助课题

必要将此法与具有通过监测业务实际 QoS 来估计其资源要求的动态机制结合使用^[3]。此外, 业务统计特征的复杂性和业务服务质量要求的多样性使得等效带宽估计问题相当复杂, 不同的解析方法和排队模型不断涌现 (如非均匀 ON/OFF 业务模型^[4]), 但是精确的解析模型在线计算往往难以满足实时 CAC 判决的要求。因此, 最近出现了许多等效带宽估计的近似方法: 如 Hiramatsu 提出了利用 BP 神经网络进行 ATM 网络接纳控制^[5]; Chang 提出了基于模糊子集理论的业务量控制方法^[6]; Ndousse 用模糊联想记忆型神经网络实现了一个改进的漏桶机制^[7]; 文献 [8] 中也提出了一种基于模糊神经网络和遗传算法的 CAC 策略。然而, 以上方法都未对多类业务间共享链路带宽资源的公平性加以考虑。

本文先简述业务的等效带宽估计方法, 而后给出网络资源分配的合作对策模型, 并从各用户偏好函数乘积最大化的角度讨论业务间共享链路资源的公平性, 提出一种基于时延带宽积的业务收益函数形式, 最后利用改进的遗传算法求解上述函数优化问题, 得到业务源的可获得带宽并对新的呼叫进行接入或拒绝判决。仿真结果证实了本文方法的有效性。

2 动态带宽分配的合作对策模型

2.1 业务的等效带宽模型

本节我们简要给出对网络接入的突发业务源估计其等效带宽的方法。由于业务到达过程的统计模型多种多样, 因而当多业务流在网路上形成统计复用时, 单个业务所占有的带宽只具有某种平均意义。设业务流的离散到达过程为 $\{X_1, X_2, \dots\}$, 服务速率为 S , 特定的观察时间内有 W 个这样的离散到达且 W 足够大, 关于其渐近分布有估计:

$$\hat{X}_1 = \sum_{k=1}^W X_k, \quad \hat{X}_2 = \sum_{k=W+1}^{2W} X_k, \quad \dots \quad (1)$$

假定到达过程平稳, 根据 (1) 式得到的部分和序列接近于独立同分布, 由大偏差理论 (Large deviation theory), 等效带宽 b 可依以下各式估计。

$$\hat{\lambda}_n^W(\theta) = \frac{1}{W} \log \frac{W}{n} \sum_{i=1}^{n/W} e^{\theta X_i}, \quad (2)$$

$$\hat{\delta}_n^W = \sup\{\theta | \hat{\lambda}_n^W(\theta) \leq S\theta\}, \quad (3)$$

$$b = \lim_{n \rightarrow \infty} \frac{\hat{\lambda}_n^W(\hat{\delta}_n^W)}{\hat{\delta}_n^W}. \quad (4)$$

根据业务的等效带宽和网络的服务速率可确定此业务的信元丢失率, 有关等效带宽估计的详细论述参见文献 [4,9]。值得指出的是, 业务的等效带宽表征了网络为确保该业务的 QoS 而必须分配的资源, 这也是后文讨论 CAC 判决准则的基础。

2.2 基于业务收益函数的合作对策模型

早期的动态资源分配策略只考虑业务的服务速率和特定连接的吞吐能力, 而对相互竞争的呼叫在接入过程中的业务类别选择则较少考虑。近来, 在 B-ISDN 的应用中, 多业务流共享网络资源的公平性已越来越引起研究者的重视^[10], 多业务共享资源的公平分配问题非常适合于用多人合作对策模型来描述。早期利用对策模型讨论网络资源分配的文章如文献

[11,12]. 文献 [11] 比较了 Raiffa、Nash 和改进的 Thomson 解的特点和适用范围, 但仅从业务阻塞概率平均化的角度给出了几种简化的带宽分配模型。

以下我们先介绍二人合作对策模型, 而后定义一种基于业务时延带宽积的收益函数, 并给出在此模型下的 CAC 判决准则, 最后我们通过遗传算法求解对策模型下的优化函数。为讨论方便, 设有两类呼叫业务竞争网络资源, 网络带宽在它们之间的分配方式构成一对策集, 对策的结果用二局中人的收益 $u = \{u_1, u_2\}$, $u_i \in R$ 来衡量, 设局中人所有可能的收益集合为 U , 并假定 U 为有界的凸闭集 (该假定只是在传统对策论模型中用来保证最优解的存在唯一性, 在用遗传算法求解过程中并未用到 U 的凸性)。注意到二人合作对策的最终结果具有以下性质: 局中人一方收益的增加必然会引起另一方收益的减少, 且每人的合作收益都超过非合作收益。设想当两类业务的呼叫请求发生冲突时, 网络端可以通过某种“协商”的办法保证它们对链路资源利用的公平性, 这一过程称为合作对策。为方便讨论, 假设 U 中的元素已规范化至 $[0,1]$, 定义局中人的偏好函数如下:

$$\nu_1 = u_1 + \beta(1 - u_2), \quad (5)$$

$$\nu_2 = u_2 + \beta(1 - u_1). \quad (6)$$

式中 β 是一加权因子, 在二人合作对策模型中, 通过求解最优化问题:

$$\max_u \{\nu_1 \cdot \nu_2\}. \quad (7)$$

可以使二局中人获得的收益在一定程度上比较“公平”。当 $\beta = 0$ 时优化问题 (7) 式的解称为 Nash 解, $\beta = 1$ 时的解称为 Raiffa 解, $\beta = -1$ 时的解称为改进的 Thomson 解。在多人合作对策模型下, 偏好函数的形式变为

$$\nu_j = u_j + |\beta(N-1)| - \beta \sum_{j \neq i} u_i. \quad (8)$$

关于各种解的经济学含义参见文献 [13]。

应用合作对策模型时, 选取合适的收益函数是保证业务间公平性的关键, 除了考虑呼叫请求端业务源之间带宽的差异, 还必须注意到该业务传输的数据量以及整个传输链路上资源利用的情况。下面我们给出的用户收益函数即同时将上述因素考虑在内, 并且用一乘性加权函数控制服务质量对用户满意程度的影响。设有如下的代价函数模型: C 为用户原始的数据包总长, C' 为加上由于传输差错或信元丢失而导致重发的数据后的总长, 将上述因素都视为信元丢失就有 $C' = C[1 + \alpha(L)]$, L 为单交换节点上的信元丢失率, $\alpha(L)$ 是由于信元丢失引起数据重发的影响因子。设 B 为链路分配的可用带宽, T 为源到目的节点的传输时间, 则考虑数据重发因素后业务的平均时延定义为

$$D(L, B) \approx C'/B + k(C'/B + 2T)nC'L/(1 - nC'L). \quad (9)$$

式中 k 为一窗宽系数, 在滑动窗管制机制下可取窗宽的一半, 在漏桶调节机制下为一与业务突发特性有关的变量, 其上界不超过漏桶容量; n 为源到目的所经过的交换节点数, 在 (9) 式中潜在假定了各节点上的信元丢失率近似相等。

基于时延带宽积的收益函数定义为

$$C(L, B) = f(L)[C' + k(C' + 2BT)nC'L/(1 - nC'L)], \quad (10)$$

式中 $f(L)$ 为一加权函数, 根据不同业务的服务质量要求, 随信元丢失率的增大而衰减。业务数据长度与链路的时延带宽积体现了传输该类业务对网络资源的要求, 因而我们将它看作是各种不同业务的收益函数, 令

$$u_i = C_i(L_i, B_i), \quad i = 1, 2, \dots, N, \quad (11)$$

则对策问题 (7) 式的解在一定程度上保证了不同业务占用网络资源的公平性。

2.3 考虑业务间公平性的连接接纳控制策略

引入合作对策模型后, 用户接入端的连接接纳控制包含以下过程: 当第 i 类业务的呼叫请求到来时, 网络按 2.1 节中的方法估算业务的等效带宽 \hat{b}_i , 如果业务的统计参数模型已知, 也可通过数值计算得到等效带宽。同时遗传算法不断搜索对策问题 (7) 式的解, 从而得到该类业务所能获得的带宽 B_i 。对于用户业务形式固定的情形, 网络端可以离线或在信令初期完成遗传搜索; 对于具有带宽重分配机制的网络, 遗传搜索在线不断进行 B_i 的调整。设已接入的第 i 类业务占用的链路带宽为 \bar{B}_i , 则网络根据 $B_i - \bar{B}_i$ 是否大于 \hat{b}_i 决定接入或拒绝该业务源。实际上, 当有 i, j 两类业务的呼叫同时竞争网络资源时, 网络将优先接入已占用带宽尚未超过其所能获得最大带宽的业务类别。这种 CAC 判决机制包含了动态的带宽预留, 与传统的在呼叫阻塞概率平均化约束下作 CAC 判决^[10]相比, 其带宽预留机制根据不同业务偏好有了更大的灵活性。下节将给出遗传算法用于求解 (7) 式的步骤, 并通过仿真说明适当选择 (5) 和 (6) 式中偏好函数的加权因子, 即可满足不同类业务间的公平性要求。

3 利用遗传算法进行呼叫接纳控制

3.1 遗传算法的基本原理

遗传算法 (Genetic Algorithm: GA) 通常把待优化问题归结为一下形式的寻优问题:

$$\min\{f(C) | C \in IB^N\}. \quad (12)$$

这里, 假设对 $\forall C \in IB^N = [0, 1]^N$, $0 < f(C) < \infty$, 显然 (7) 式满足上述要求。利用 GA 解决此类优化问题的主要步骤如下: 首先要将待解问题的候选解按照一定的规则编码成字符串 (通常是 0, 1 组成), 一条字符串则类似于一条基因串, 其中蕴含着该问题解的信息。然后, 再给出评价各字符串 (基因串) 性能的办法, 通常是定义一适合度函数 f , 以评价基因串的优劣。GA 在开始时随机地产生一定数目的个体, 通过对基因串进行组合操作, 得到新的个体群; 而后计算其中各基因串的适合度函数值 $f_i (i = 1, 2, \dots)$, 得到对这些基因串优劣的评价; 再根据优劣分配繁殖机会, 优的个体分配的机会多, 产生的子代也就多; 劣的个体分配的机会少, 产生的子代少, 这样搜索过程按照一代一代逐步朝着更好的解的方向进化。下面即为一种常规的 GA 的算法格式:

```

产生初始的个体群
计算每个个体的适合度函数值
进行选择
REPEAT
    进行交配
    进行变异
    计算每个个体的适合度函数值

```

进行选择

UNTIL 满足终止条件

GA 的交配和变异等各种算子在不同的优化问题中可能有不同的实现方法,其效率也差别很大,具体论述参见文献 [14,15]。

3.2 连接接纳控制的数值仿真结果

上节中我们介绍了遗传算法 (GA) 的基本原理,为实时、有效地求解优化问题 (7) 式,我们采用文献 [14,15] 中改进交配和变异算子之后的 GA。GA 待优化的权向量为各类业务所能获得的带宽 $B_i (i=1,2,\dots,N)$, 当第 i 类业务源已接入 n_i 个业务时,该类业务在此交换节点处的信元丢失率可根据其等效带宽进行估计^[9]。将 (10), (11) 式规范化并代入 (7) 式即得 GA 的适合度函数,仿真实验表明,采用改进的遗传算法通常只须 20 次左右繁衍即可得到满意的结果。

仿真过程如下:设有两类突发业务源,第一类业务的峰值位率为 64kbps,平均位率为 22kbps,平均突发长度为 100 信元,整个路由途经 4 个交换节点,传输时延为 4ms, QoS 要求为 10^{-4} ;第二类业务的峰值位率为 10Mbps,平均位率为 1Mbps,平均突发长度为 300 信元,途经 2 个交换节点,传输时延为 0.28ms, QoS 要求为 10^{-8} 。设业务呼叫端的链路容量为 155.5Mbps, (10) 式中的加权函数取为如下经验公式:

$$f(L) = e^{-10(L-QoS)}, \quad \alpha(L) = 1.05L. \quad (13)$$

遗传算法对规范化的分配带宽编码长度为 15 位。经计算机仿真,在不同的缓冲容量和两类业务不同的发送数据长度之下得到的 CAC 判决边界如表 1 所示。

表 1 两类业务在不同加权因子和不同发送数据量之下的 CAC 判决边界

加权因子 β	缓冲容量 (信元)	第一类业务 数据长度	第二类业务 数据长度	接入第一类 业务个数	接入第二类 业务个数
Nash 0	10^3	10^5	10^7	4046	3
Raiffa 1	10^3	10^5	10^7	4748	1
Thomson-1	10^3	10^5	10^7	2430	8
-0.25	10^3	10^5	10^7	3460	4
-0.75	10^3	10^5	10^7	2729	7
0.25	10^3	10^5	10^7	4328	2
0.75	10^3	10^5	10^7	4625	1
Nash	10^3	10^5	10^8	2628	7
Raiffa	10^3	10^5	10^8	3296	3
Thomson	10^3	10^5	10^8	1368	13
Nash	10^3	10^3	10^8	1966	12
Raiffa	10^3	10^3	10^8	4338	4
Thomson	10^3	10^3	10^8	76	17
Nash	10^2	10^5	10^7	2495	3
Raiffa	10^2	10^5	10^7	3868	1
Thomson	10^2	10^5	10^7	1245	5

由表 1 可以看出,发送数据量大的业务类别从网络链路得到的带宽资源相对较大,由于第二类业务相对于第一类为宽带业务,如果不考虑业务间的公平性,当第一类业务的接入呼叫达到一定数量后,网络的剩余带宽将长期处于不够分配给第二类业务呼叫的状态。而合作对策模型的解具有以下性质:保证为宽带业务留有可接入的带宽,而且业务传输路径越长、发送数据量越大,网络为其保留的带宽资源也越大。由表 1 还可看出,同样的业务源在同样

的偏好函数加权因子下, 当网络的缓冲容量减小时得到 CAC 边界会有所不同, 这是由于业务的等效带宽变大。根据 2.3 节给出的 CAC 判决准则, 显然对策模型的解域不会超出不考虑业务间公平性时的 CAC 边界。

偏好函数加权因子的选择对不同类业务间公平性的影响很大, 当 β 由 -1 逐渐增至 1 时, 第一类业务获得的带宽也逐渐增大。 $\beta = 1$ 时偏好函数考虑自己收益的同时也考虑对方的损失; $\beta = -1$ 考虑的是局中人合作后的共同收益; $\beta = 0$ 时局中人只考虑自己的收益; 这三者都代表一定的业务公平性指标。在仿真测试中我们发现, 以 Nash 解作为 CAC 判决的一个附加约束, 具有较好的性能。当 $\beta = 0$ 且传输数据长度发生变化时, 第一类业务获得链路带宽的上界为 $28\%—87\%$, 第二类业务获得链路带宽的上界为 $46\%—78\%$, 说明动态带宽预留机制根据业务数据长度的变化可灵活调整 CAC 判决边界, 如果我们将 CAC 同业务量管制和缓冲调度策略结合使用, 可望进一步提高链路利用率。

4 结 论

本文先简述了基于动态带宽分配的连接接纳控制原理, 而后从合作对策模型的角度讨论业务间共享链路资源的公平性问题, 提出了一种基于时延带宽积的业务收益函数形式, 并通过遗传算法对待优化的对策函数进行寻优, 以决定对业务呼叫请求的接入或拒绝, 此方法能够保证不同带宽和服务质量要求的业务共享网络资源的公平性。进一步的研究工作包括建立基于 QoS 的网络传输费用模型, 分析缓冲资源分配和流量平滑机制对本文 CAC 策略的影响, 以及比较传统流量控制策略和基于对策模型的控制策略采用真实业务源在 ATM 网络上的传输性能。

参 考 文 献

- [1] Hsu I, Walrand J. Dynamic bandwidth allocation for ATM switches. to appear in Journal of Applied Probability, 1996, also available via <http://www.path.berkeley.edu/~wlr/>.
- [2] Perros H G, Elsayed K M. Call admission control schemes: A review. IEEE Communication Magazine, 1996, 34(11): 82-91.
- [3] Chong S, Li S Q, Ghosh J. Predictive dynamic bandwidth allocation for efficient transport of real-time VBR video over ATM. IEEE J. of Selected Areas in Communications, 1995, 13(1): 12-23.
- [4] Lee H W, Mark J W. Capacity allocation in statistical multiplexing of ATM sources. IEEE/ACM Trans. on Networking, 1995, 3(2): 139-151.
- [5] Hiramatsu A. Integration of ATM call admission control and link capacity control by distributed neural networks. IEEE J. of Selected Areas in Communications, 1991, 9(7): 1131-1137.
- [6] Chang C J, Cheng R G. Traffic control in an ATM network using fuzzy set theory. Proc. IEEE INFOCOM'94, Toronto, Canada: 1994, 1200-1207.
- [7] Ndousse T D. Fuzzy neural control of voice cells in ATM networks. IEEE J. of Selected Areas in Communications, 1994, 12(9): 1488-1494.
- [8] 张良杰, 李衍达. ATM 网络中基于模糊神经网络与遗传算法的动态带宽分配技术. 北京: 清华大学自动化系网络与信息研究组技术报告, 1995 年 12 月.
- [9] 陈惠民. ATM 网络业务流量综合控制的初步研究. 北京: 清华大学自动化系毕业设计论文, 1996 年 6 月.
- [10] Bolla R, Davoli F, Marchese M. Bandwidth allocation and admission control in ATM networks with service separation. IEEE Communication Magazine, 1997, 35(5): 130-137.
- [11] Dziong Z, Mason L G. Fair-efficient call admission control policies for broadband networks—A game theoretic framework. IEEE/ACM Trans. on Networking, 1996, 4(1): 123-136.

- [12] Sairamesh J, Ferguson D, Yemini Y. An approach to pricing, optimal allocations, and quality of service. Proc. of the INFOCOM'95, Boston, USA: 1995, also available via <http://www.ics.forth.gr/~ramesh/pricing.html>.
- [13] Shubik M. Game theory in the social science: Concepts and solutions. The MIT Press, 1982. 336-368.
- [14] Zhang L J, Mao Z H, Li Y D. An improved genetic algorithm based on combinative theory and fuzzy reasoning and its applications. International Conference on Neural Information Processing, Korea: 1994, 180-185.
- [15] Zhang L J, Mao Z H, Li Y D. Mathematical analysis of crossover operator in genetic algorithms and its improved strategy. IEEE International Conference on Evolutionary Computing, Australia: 1995, 412- 417.

A CONNECTION ADMISSION CONTROL SCHEME BASED ON THEORETIC GAME MODEL IN ATM NETWORKS

Chen Huimin Wang Pu Li Yanda

(Department of Automation, Tsinghua University, Beijing 100084)

Abstract In this paper, the main schemes of connection admission control (CAC) in ATM networks are briefly overviewed especially the principle of dynamic bandwidth allocation. Then the fair share of the bandwidth among different traffic sources is analyzed based on cooperative game model. A CAC scheme combining genetic algorithm (GA) to optimize the bandwidth-delay-product formed utilization function that ensures the fair share of accepting/rejecting the incoming calls is proposed. Simulation results show that our scheme ensures fairness of the shared bandwidth to different traffic sources.

Key words Connection admission control, Cooperative game model, Genetic algorithm, Dynamic bandwidth allocation, ATM network

陈惠民: 男, 1972年生, 硕士生, 现从事 ATM 网络的业务建模、流量控制和分布式计费管理策略研究。

王 普: 男, 1945年生, 副教授, 从事多媒体信息网络、智能家电产品的技术及应用研究。

李衍达: 男, 1936年生, 教授, 中国科学院院士, 从事信号处理和计算智能理论、方法及应用的研究, 包括综合利用多种信息处理的理论及方法(如小波变换、分形信号处理)用于油气勘探、开发, 利用人工神经网络、模糊理论、进化算法等技术于信息网络的智能控制、网络信息的发掘及提取以及工业生产的过程控制等。