

NDPS: 一种无线多媒体网络分组调度算法

余 荣, 贾志鹏, 梅顺良

(清华大学电子工程系微波与数字通信国家重点实验室, 北京 100084)

摘要: 分组调度是实现未来无线多媒体网络的关键技术之一。解决该技术的主要困难在于无线链路的高差错率、业务类型的多样性和分组到达模型的未知性。该文引入马尔可夫决策过程对分组调度过程进行建模, 运用神经动态规划方法求解相应的马尔可夫决策过程问题。提出一种无线多媒体网络分组调度算法(NDPS)可以同时实现3个性能目标: 对不同业务类型提供差分服务, 最大化无线带宽的利用率和保证服务公平性。仿真实验结果证明, NDPS算法比两种流行的调度算法具有更好的性能。

关键词: 分组调度; 多媒体网络; 马尔可夫决策过程; 神经动态规划

NDPS: Packet Scheduling Algorithm for Wireless Multimedia Networks

YU Rong, JIA Zhi-peng, MEI Shun-liang

(State Key Laboratory on Microwave & Digital Communications, Department of Electronic Engineering, Tsinghua University, Beijing 100084)

【Abstract】 Packet scheduling technology is one of the important technical issues in designing wireless multimedia networks. The key difficulties lie in the high variability of wireless channel capacity and the unknown model of packet arrival processes. This paper views the packet scheduling problem as a Semi-Markov Decision Process(SMDP), and approximately solves the SMDP problem by the methodology of Neuro-Dynamic Programming(NDP). The proposed algorithm, called Neuro-Dynamic Programming Scheduling(NDPS), can simultaneously achieve three performance objectives: QoS differentiation and guarantee, high bandwidth utilization, and both short-term and long-term fairness. Simulation experimental result demonstrates that NDPS outperforms two popular scheduling algorithms.

【Key words】 packet scheduling; multimedia networks; Markov decision processes; neuro-dynamic programming

1 概述

通信网络中的分组调度算法负责管理信道带宽的分配和复用, 对整个通信网络的性能好坏有至关重要的影响。本文主要研究无线多媒体网络中的调度算法。与有线网络调度算法相比, 无线多媒体网络调度算法的设计有2个难点: (1)无线信道中信息传递高差错率。具体来说, 由于受到无线传播多径和衰落等因素的影响, 各个用户和基站之间的无线信道的容量是随机、异步变化的。(2)业务类型的多样性和分组到达过程模型的不确定性。在研究传统调度算法时, 通常假设到达过程服从泊松(Poisson)分布。然而在无线多媒体网络中这种假设不成立, 因为实际的分组到达过程是未知的。

无线多媒体网络的分组调度算法需实现3个目标: (1)对不同数据类型的分组提供差分服务。无线多媒体网络所承载的业务类型是多种多样的, 包括语音、视频、文本、图像、FTP数据、Web数据和即时短消息等类型。不同业务类型对服务质量的要求可能相差很大, 调度算法必须根据各种业务的实际需求提供服务质量保证。(2)最大化无线带宽的利用率。无线通信网络的资源之一就是无线信道带宽。调度算法应该充分利用有限的带宽传输尽可能多的信息。(3)保证服务的公平性。调度算法必须公平地对网络中的多个用户提供服务, 让用户公平地共享无线资源。

现有的分组调度算法可大致分为3类: 有线调度算法, 如先到先服务(FCFS)和轮询算法(round robin)等; 基于两状态信道模型的无线调度算法, 如CIF-Q^[1]和IWFQ等; 基于多状态信道模型的无线调度算法, 如M-LWDF等。本文采用马尔可夫决策过程理论对分组调度过程进行建模。通常来说, 求

解马尔可夫决策过程问题要求系统的状态转移模型为已知。如前所述, 在无线多媒体网络的分组调度问题中, 分组到达过程模型和无线链路状况的变化规律对于调度器来说是不可知的。所以, 基于模型的求解方法不可行, 本文引入神经动态规划方法来求解。该方法的优点在于使用者不需要完全了解系统的状态转移模型也可以通过在线学习的方式逐步逼近问题的最优解。

2 马尔可夫决策过程建模

考虑共享信道的无线蜂窝网络模型, 每个小区内的基站对该小区的用户服务, 基站负责调度用户业务流。每个用户和基站之间可以建立多个业务流连接, 各用户和基站之间的无线通信链路相互独立, 构成如图1所示的分组调度模型。假设系统中引入了链路状态监督机制(LSM), 调度器掌握所有链路的状态信息, 如瞬时可用速率等。

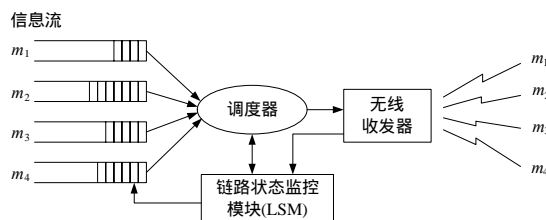


图1 分组调度系统模型

作者简介: 余 荣(1979 -), 男, 博士研究生, 主研方向: 无线传感器网络, 宽带无线接入网; 贾志鹏, 硕士研究生; 梅顺良, 教授、博士生导师

收稿日期: 2007-06-30 **E-mail:** yurong@mails.tsinghua.edu.cn

对于无线信道模型，采用图 2 所示的多状态马尔可夫模型^[2]，其中， p_{ij} 表示由状态*i*到状态*j*的转移概率。可以看到，信道状态转移仅仅发生在两个相邻状态之间。对于给定的系统误码率要求，每一个状态都对应一个最高的传输速率。

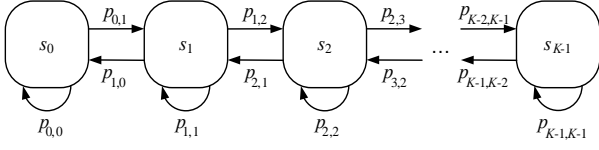


图 2 多状态信道模型

由于分组调度过程具有多阶段性和无记忆特性，因此可以将其建模为半马尔可夫决策过程(Semi-Markov Decision Process, SMDP)，具体如下：将分组调度过程按照随机事件 ω 划分为多个“阶段”， ω 表示分组到达或离开事件。系统的“状态” x 是指队列缓存中所有分组的等待时间、各业务流的信道可用速率及其实际服务速率。每当事件 ω 发生，调度器就需要做出一个调度“决策” u ，指定某一个业务流(队列)进行传输。系统“时间平均代价”的定义如下：

$$v = \lim_{N \rightarrow \infty} \frac{1}{t_N} \sum_{i=0}^{N-1} g(x_i, \omega_i, u_i) \quad (1)$$

其中， t_i 表示事件 ω_i 发生的时刻； $g(x_i, \omega_i, u_i)$ 为第 i 阶段的“代价函数”，其具体定义如下：

$$g(x_i, \omega_i, u_i) = \sum_{m=1}^M W_m F_m(r_m) \tau_m^1 I_{(u_i=m)} \quad (2)$$

其中， M 表示业务流个数； I_A 为指示函数(当事件 A 发生时， $I_A=1$ ；否则， $I_A=0$)； τ_m^1 表示第 m 个业务流首个分组的等待时间； r_m 表示第 m 个业务流的实际服务速率；常数 $W_m > 0$ ，称为优先权因子； F_m 称为公平性因子，定义为 $F_m(r_m) = (r_m/\hat{r}_m)^{-n_m}$ (常数 $n_m > 0$ ， \hat{r}_m 为第 m 个分组的标称服务速率)。可以看出，代价函数的定义同时考虑了分组调度算法的 3 个性能目标。

Bellman 最优方程是马尔可夫决策过程理论的核心，针对分组调度问题形式如下^[3]：

$$v^* E_{\Delta t} \{ \Delta t \{ x \} + h^*(x) \} = E_{\omega} \left\{ \min_{u \in U(x, \omega)} [g(x, \omega, u) + h^*(y)] \right\} \quad (3)$$

其中， $h^*(\hat{x})=0$ ； E_A 表示对随机变量 A 求数学期望； x 为当前状态； y 为下一个阶段的状态； Δt 为当前阶段的持续时间； $U(x, \omega)$ 表示备选的决策集合；称变量 v^* 为“最优平均代价”；函数 $h^*(\cdot)$ 为“最优偏差代价函数”^[3]；状态 \hat{x} 为偏差代价为零的状态，对应于系统中所有队列均为空的状态。

如果可以得到函数 $h^*(\cdot)$ 和变量 v^* 的值，就可以在每个阶段选取合适的决策 u ，使得式(3)的右边取得最小值，从而得到最优的调度策略。可见，解决分组调度所对应的马尔可夫决策过程问题的关键就在于求出函数 $h^*(\cdot)$ 和变量 v^* 。

3 神经动态规划方法求解

由于分组调度问题的状态空间(也即所有可能状态的集合)较大，精确求解函数 $h^*(\cdot)$ 和变量 v^* 的计算复杂度极高，因此本文采用神经动态规划方法(Neuro-Dynamic Programming, NDP)求得其近似解。基本原理是选取一个特定结构的参数化函数 $\tilde{h}(\cdot)$ 来近似的取代 $h^*(\cdot)$ ，并采用机器学习(machine learning)方法不断地在线更新逼近函数的参数值，使其越来越接近于最优函数 $h^*(\cdot)$ 。考虑到神经动态规划方法是所提出

的调度算法的核心，因此，将该算法命名为 NDPS(NDP scheduling)。

为了降低计算复杂度并使算法易于实现，采用了基于特征线性函数逼近结构^[3]来构造近似函数 $\tilde{h}(\cdot)$ ，其结构表示如下：

$$\tilde{h}(x, \theta) = \theta^T f(x) \quad (4)$$

其中， θ 为可调整的参数向量；函数 $f(x)$ 称为特征向量，用以描述系统状态的特征信息，并使其本身的线性组合可以准确地逼近最优函数。特征向量的每一个分量均相对应于系统中一个数据分组的特征信息。具体来说，对第 m 个业务流的第 k 个分组，定义如下的特征分量：

$$\xi_m^k = W_m F_m(r_m) \left(\tau_m^k + \frac{kl_m}{r_m} \right) \quad (5)$$

其中， l_m 表示第 m 个业务流的分组长度。可以看出式(5)的定义与式(2)的代价函数定义很接近，所不同的是多了 kl_m/r_m ，表示对第 m 个业务流的第 k 个分组的剩余等待时间的估计值。根据式(5)将式(4)展开，得到式(6)：

$$\tilde{h}(x, \theta) = \sum_{m=1}^M \sum_{k=1}^{K_m} \theta_m^k \xi_m^k \quad (6)$$

其中， K_m 表示第 m 个业务流的缓冲区大小(以分组为单位)； θ_m^k 为对应于特征分量 ξ_m^k 的待调整参数。

确定逼近函数的结构后，采用时间差分学习算法(temporal-difference learning)^[3]来实现在线学习和参数调整。时间差分学习算法的基本思想是：随着分组调度过程的进行，调度器将观察到越来越多的系统状态转移事件，逐渐学习到分组的到达特性和用户信道状况的变化规律，以此作为依据，不断地更新待调整的参数，使逼近函数越来越接近于目标函数。算法首先给参数向量 θ 和平均代价的近似值 \tilde{v} 分别设定初始值 $\theta = \theta_0$ ， $\tilde{v} = \tilde{v}_0$ ；在每一个阶段，都会更新一次 θ 和 \tilde{v} 。具体做法是首先根据当前的状态 x 、参数向量 θ 和变量 \tilde{v} 做出一个贪婪(greedy)决策：即选取最合适的决策 u ，使得 $g(x, \omega, u) + \tilde{h}(\tilde{y}, \theta)$ 的取值最小。其中， $\tilde{\omega}$ 和 \tilde{y} 分别代表对下一个随机事件 ω 和下一阶段系统状态 y 的估计。可以简单地假设在当前的分组传输结束之前没有新的分组到达，这样就很容易计算上述两项估计值。在第 i 阶段，参数向量 θ 和近似平均代价 \tilde{v} 的更新公式分别为

$$\theta_i = \theta_{i-1} + \gamma_i d_i \nabla_{\theta} \tilde{h}(x_{i-1}, \theta_{i-1}) \quad (7)$$

$$\tilde{v}_i = \tilde{v}_{i-1} + \eta_i [g(x_{i-1}, \omega_{i-1}, u_{i-1}) - \tilde{v}_{i-1} \Delta t_i] \quad (8)$$

其中， γ_i 和 η_i 是迭代步长； d_i 即是所谓的时间差分(temporal difference)，其计算方法如下：

$$d_i = g(x_i, \omega_i, u_i) - \tilde{v}_i \Delta t_i + \tilde{h}(x_{i+1}, \theta_i) - \tilde{h}(x_i, \theta_i) \quad (9)$$

值得注意的是，由于采用了式(4)的线性逼近结构，式(7)中的求梯度运算 $\nabla_{\theta} \tilde{h}(x, \theta) = f(x)$ ，大大简化了运算。

NDPS 的工作流程如图 3 所示，主要包括 5 个步骤：
(1)根据式(6)计算在上一阶段系统的代价 g ，并更新各个业务流的实际服务速率；(2)观察当前的状态 x ，运用时间差分学习方法更新 θ 和 \tilde{v} ，并得到新的函数 $\tilde{h}(\cdot, \theta)$ ，这一步将用到式(7)和式(8)；(3)判断随机事件 ω 是否是分组离开(ω 可能代表分组离开或分组到达)；(4)若事件 ω 为分组离开，说明上一次的传输结束，需要选择一个新的业务流传输，可根据贪婪算法作出决策；若事件 ω 是分组到达，还需根据信道是否空闲来决定是否调度新业务流传输；(5)若贪婪算法找到适合

传输的业务流, 该业务流将在下阶段传输, 若所有的业务流都不可用, 导致贪婪算法得出空集, NDPS 将无法发出调度指令, 本阶段的任务结束。

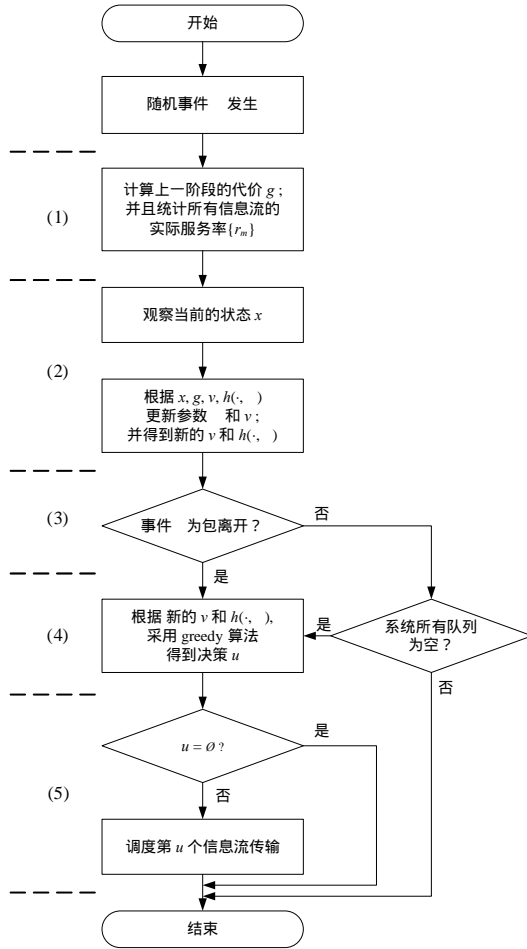


图3 NDPS 的工作流程

4 仿真实验和结果

本节将NDPS算法和2个流行的调度算法:CSDP-WRR^[4]和CIF-Q^[1]进行比较。如表1所示,考虑5个业务流的调度系统,所有的业务流均严格按照文献[5]的参考模型产生。假设仅有业务流数据-2和数据-3受链路错误的影响,其链路状况用六状态的信道模型描述(见图2)。信道状态s₁~s₆的传输速率分别设定为全速率2 Mb/s的100%,80%,60%,40%,20%和0%。在链路模式1和模式2下,各个信道状态的稳态概率分布分别为 $\pi_1 = (0.52, 0.10, 0.10, 0.10, 0.10, 0.08)$ 和 $\pi_2 = (0.17, 0.17, 0.17, 0.17, 0.15)$ 。在总共1200s的仿真过程中,假设链路差错只发生在前400s,剩余的无差错时间用以评估算法的长时公平性。式(2)中相关参数的具体取值为: $n_{\{1,2\}} = 1$, $n_{\{3,4,5\}} = 3$, $W_{\{1,2\}} = 1$, $W_{\{3,4,5\}} = 10^{-2}$ 。

表1 仿真实验中的业务流参数

业务流	语音	视频	数据1	数据2	数据3
业务模型	IDP	2IRP	4IPP	4IPP	4IPP
分组长度/B	66	188	192	192	192
平均速率/(Kb·s ⁻¹)	22.4	0.19	0.8	0.8	0.8
峰值速率/(Kb·s ⁻¹)	64	0.4	1.8	1.8	1.8
链路状况	好	好	好	模式1	模式2

仿真结果如图4、图5所示。从图4中可见,相比于其他两种调度算法,NDPS不仅可以为实时业务流提供更好的服务质量,如更低的平均和最大延时,见图4(a)、图4(b),

还同时提高了非实时业务流的吞吐量,见图4(c)。这说明NDPS在调度性能上有了全面的提高。如图5所示,3个数据类型业务流的服务曲线在仿真的前400s内是逐渐发散的,而在后800s内则又慢慢的收敛在一起。这表明NDPS同时实现了短时和长时公平性。

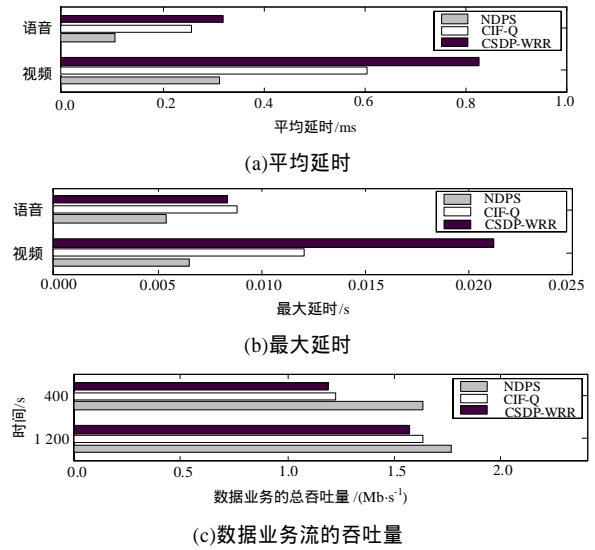


图4 3种调度算法的性能比较

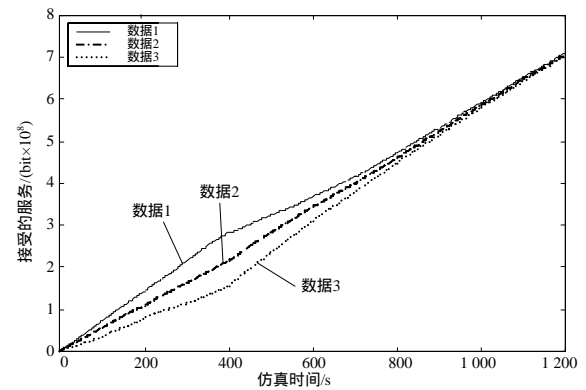


图5 数据业务的服务曲线

具体地讲,它保证了无链路差错的业务流的短时公平性(在短时间内得到更多的服务)和受链路差错影响的业务流的长时公平性(在足够长的时间内,必将逐渐得到补偿服务),这与文献[1]中的长时和短时公平性定义完全吻合。

参考文献

- [1] Ng T S E, Stoica I. Packet Fair Queuing Algorithms for Wireless Networks with Location-dependent Errors[C]//Proc. of INFOCOM'98. San Francisco, CA, USA: [s. n.], 1998: 1103-1111.
- [2] Wang Hongshen, Moayeri N. Finite-state Markov Channel A Useful Model for Radio Communication Channels[J]. IEEE Transactions on Vehicular Technology, 1995, 44(1): 163-171.
- [3] Bertsekas D P, Tsitsiklis J N. Neuro-dynamic Programming[M]. Belmont, MA: Athena Scientific Publisher, 1996: 512.
- [4] Bhagwat P, Krishna A, Tripathi S. Enhancing Throughput over Wireless LANs Using Channel State Dependent Packet Scheduling[C]//Proc. of INFOCOM'96. San Francisco, CA, USA: [s. n.], 1996: 1133-1140.
- [5] Baugh C R, Huang J. Traffic Model for 802.16 TG3 MAC/PHY Simulations[Z]. (2001-03-02). <http://wirelessman.org/tg3/contrib/802163c-01-30r1.pdf>.