

用于可变比特率视频通信量预测的 自适应训练及删剪算法*

熊 涛 常胜江[†] 申金媛 张延

(南开大学现代光学研究所, 光电信息技术科学教育部重点实验室, 天津 300071)

(2003 年 10 月 8 日收到, 2004 年 11 月 25 日收到修改稿)

提出一种自适应神经网络模型对可变比特率视频通信量进行非线性自适应预测, 并采用基于递归最小方差的自适应学习及删剪算法对抽头延迟神经网络进行训练和结构优化. 仿真实验表明, 该模型能够实现对复杂视频通信量序列的高精度预测, 满足实时快速的预测要求.

关键词: 视频通信量预测, 神经网络, 递归最小方差, 删剪算法

PACC: 8730

1. 引 言

目前, 数字视频^[1]在许多领域已有着广泛的应用. 数字电视和高清晰度电视的发展、可视电话的日益普及均说明了这一点. 在图像/视频压缩方面取得的巨大成功使得数字视频的有效存储与低速率传输已成为可能, 由此产生了可变比特率(variable bit rate, 简记为 VBR) 视频的概念. 正是由于数字视频技术的快速发展, 使得 VBR 视频业务在基于异步传输模式的宽带综合业务数据网中占据大部分的带宽. 因此在现有数字通信技术的带宽瓶颈下, 解决 VBR 视频业务的高服务质量传输问题, 是构建未来信息高速公路的一个关键环节. 解决这一问题有三个基本途径: 一是加大网络信道自身的带宽, 二是压缩视频信息的流量, 三是改善信道带宽的利用效率, 三者相辅相成, 缺一不可. 本文侧重于研究第三个方面即改善信道带宽的利用效率问题.

异步传输模式网络是一个基于连接的通信网络, 若要建立一个连接, 就需要为这个新的连接分配所需要的带宽资源. 静态分配方法是一次到位的带宽分配方法, 由于 VBR 视频通信量具有时变、非线性、突发性等特征, 静态方法很难分配合适的带宽资

源, 为保证高质量服务, 网络资源利用效率很低. 为提高网络带宽在多连接条件下的利用效率, 就要求动态分配带宽, 从而合理地利用资源, 提高网络的传输效率. 由于 VBR 信源统计特性的复杂性, 依赖有明确数学关系的算法解决带宽的动态分配是非常困难的, 最有效的一类方法是通过通信量特征在线测量与学习, 预测未来的带宽要求, 从而根据预测值动态分配带宽. 因此, 高精度的预测方法, 不仅能够视频分配合适的带宽, 达到高质量、低延迟抖动的视频服务, 而且提高了网络的利用效率. 线性预测方法^[2, 3]是其中最常用的方法, 由于它的计算复杂度较低, 适用于在线学习方式, 即预测器的参数在预测过程中根据当前预测误差进行更新, 但用线性模型预测非线性的 VBR 视频通信量在理论上就存在不足, 它的预测精度不高.

人工神经网络是一种仿脑信息处理模型, 它可以被用来解决复杂的现实问题^[4, 5]. 由于它具有不依赖先验知识或规则为前提的自适应学习能力, 因而在那些实时性要求很高且无法用确切数学模型描述的智能信息处理问题的求解上, 神经网络方法具有其他方法无可比拟的优势, 人们已经尝试使用神经网络方法^[6-11]来解决 VBR 视频通信量的预测. 文献 [6] 采用多层前馈神经网络对随机模拟产生的自回

* 天津市自然科学基金重点项目(批准号: 023800811) 国家自然科学基金(批准号: 60277022, 60477009) 教育部博士点基金(批准号: 20030055022) 和南开大学科技创新基金资助的课题.

[†] 通讯联系人. E-mail: sjchang@nankai.edu.cn

归型视频通信量进行预测,其输出神经元采用非线性激活函数.文献[7]已经分析了非线性的输出激活函数并不适合作时间序列预测,因此这类前馈神经网络并不能给出精确的预测.文献[8]采用输出为线性激活函数的抽头延迟神经网络(tapped delay neural network,简记为TDNN)对运动图像专家组(moving picture experts group,简记为MPEG)信源通信量进行了预测,预测精度有所提高,但由于网络训练完成之后其权重参数不再发生改变,属于静态预测模型.而实际的VBR视频通信量序列的统计性质是时变的,用静态模型实现长时间的VBR视频通信量的高精度预测是非常困难的,必须采用有在线学习能力的预测模型.文献[9,10]采用的管线循环神经网络(pipelined recurrent neural network,简记为PRNN)模型对于VBR视频通信量能给出较高的预测精度,但由于PRNN采用实时循环学习算法来在线修改权重,要求在视频采样间隔内完成实时循环学习的多次迭代,计算的复杂度高,实时应用存在困难.文献[11]提出了一种基于TDNN的自适应神经网络模型,该模型分为离线学习、在线学习两个阶段,离线学习主要是为了获得先验知识,在线学习使得网络能不断地适应当前通信量的变化.该模型要求权重的调整尽可能对先验知识的影响最小,对MPEG通信量的预测得到了令人满意的结果,但是该模型需要一个大规模的矩阵去记忆先验知识,先验知识的改变用一个与该矩阵有关的能量函数表示,优化该能量函数的计算复杂度非常高,因此该模型的实时性能很差,实际应用存在困难.

本文提出了一种自适应神经网络模型,网络的训练分为离线学习和在线学习两个阶段.离线阶段网络训练采用基于递归最小方差(recursive least square,简记为RLS)的自适应学习及删剪算法以获取先验知识,同时通过删除不重要的权重,以删除相应的冗余隐层神经元,从而获得最佳的网络拓扑结构,提高网络的推广能力,降低网络在线学习的计算复杂度.网络离线学习结束后,在预测过程中,网络根据当前的预测误差,仍然采用RLS算法进行在线学习,使网络能够在兼顾先验知识的前提下适应当前通信量的变化情况,提高网络的预测精度.由于神经网络经离线学习后具有了先验知识,预测过程中的在线学习的收敛速度变得很快,能满足实时快速的预测要求.RLS算法是二阶学习算法,能在学习过程中自动估算每一次的学习步长,因此网络的收敛

速度很快,能够实现对突发度较高的VBR视频通信量的高精度预测.

2. VBR 视频通信量序列的非线性自回归模型

VBR视频是根据视频序列图像本身的信息量冗余,经过各种图像/视频压缩算法压缩后所产生的图像帧序列,其图像帧中的数据量就是VBR视频通信量.图像/视频的压缩编码已被制定出几个国际标准,如静止图像压缩标准联合图像专家组(Joint Photographic Experts Group,简记为JPEG),电话/电视会议的视频压缩标准H261, H263,活动图像压缩标准MPEG-1, MPEG-2, MPEG-4等.为了对VBR视频进行合理的带宽分配,需要对VBR视频信源进行建模,人们已建立了VBR视频通信量序列的各种数学模型^[12-14].利用前馈神经网络进行通信量预测是基于非线性自回归模型,即

$$x(n) = f(x(n-1), \dots, x(n-p)) + e(n), \quad (1)$$

式中, $\{x(n)\}$ 表示通信量时间序列, $f(\cdot)$ 是一非线性函数, $\{e(n)\}$ 是 n 时刻的系统误差,一般认为 $\{e(n)\}$ 是一个均值为零、独立、有确定分布的随机过程.值得指出的是, MPEG等压缩算法有不同的编码模式,如帧内(I)预测(P)与双向(B)模式,则(1)式中的 $\{x(n)\}$ 表示由某种特定编码模式产生的子序列.

(1)式表明,当前时刻流量 $\{x(n)\}$ 是过去流量 $x(n-1), x(n-2), \dots$ 的一个函数.因此,如果函数 $f(\cdot)$ 已知,就可以用过去的通信量预测将来的通信量,但困难在于(1)式中的非线性函数是未知的,这一函数的映射过程可以通过TDNN模型的学习得到.

3. 神经网络模型及算法

虽然静态TDNN模型能够映射出一个非线性函数关系,但其函数形式却是静态的.实际上,即使是同一VBR视频业务,其统计性质也会随时间发生变化,表现出非平稳特性,即(1)式中的函数 $f(\cdot)$ 是个随时间变化的函数,所以静态TDNN方法的预测效果并不理想.针对以上问题,本文把TDNN的训练分为离线和在线两个阶段,并采用RLS算法训练网络,使其具有自适应预测的能力,满足实时快速的预测要求.

3.1. TDNN 的拓扑结构和符号说明

TDNN 模型的拓扑结构如图 1 所示,它是一个三层前馈神经网络,与一般的前馈网络模型不同的是,TDNN 模型的输出神经元采用线性激活函数,网络的输入是经过延迟的时间序列值.其输入、输出函数关系为

$$\hat{x}(n) = \sum_i w_i^2 f(w_i^1 X(n-1)) - \theta. \quad (2)$$

对照 TDNN 模型,为了方便起见,我们给出本文所用符号的意义: $X(n-1)$ 为第 $(n-1)$ 时刻的输入矢量, θ 为输出神经元的偏置, N_n 为第 n 层神经元总数(包括阈值单元), N_w 为网络权重总数, $w_{i,j}^n(t)$ 为神经元 (n, j) 与第 n 层第 j 个神经元与 $(n+1, i)$ 之间的互连权重,其中

$$X(n-1) = [x(n-1) \dots x(n-p)]^T, \quad (3)$$

$$w^n = [(w_1^n)^T, (w_2^n)^T, \dots, (w_{N_{n+1}-1}^n)^T]^T, \quad (4)$$

$$w_i^n = [w_{i,1}^n, w_{i,2}^n, \dots, w_{i,N_n}^n]^T. \quad (5)$$

可以把所有的权重排列成一维列矢量的形式

$$w = [(w^1)^T, (w^2)^T]^T. \quad (6)$$

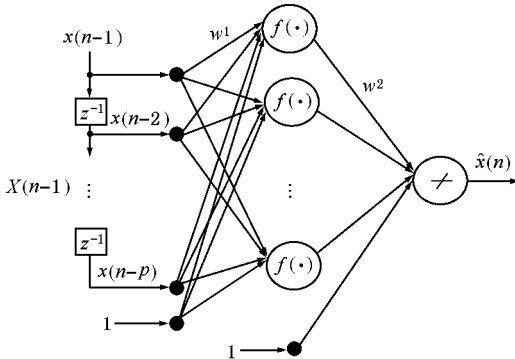


图 1 TDNN 网络结构

3.2. RLS 自适应训练算法

RLS 算法源于线性时间序列分析与预测^[15],对于非线性时间序列,通过一级近似,RLS 算法同样适用.互连权重矢量 w 被看作一个稳定非线性动力学系统的状态,假定第 n 个训练模式送入网络,对于 TDNN 模型,系统应满足下列静态方程:

$$u(n) = u(n-1) = w_0, \quad (7)$$

$$d(n) = h(u(n)) + \epsilon(n). \quad (8)$$

估算的权重 $\hat{u}(n)$ 应使下列误差函数为最小:

$$\xi(n) = \sum_{j=1}^n |d(j) - h(u(n))|^2 \lambda^{n-j}. \quad (9)$$

这里 λ 是遗忘因子,满足 $0 < \lambda < 1$ 且接近于 1.可将 $h(u(n))$ 在 $\hat{u}(n-1)$ 附近做泰勒展开,得到

$$h(u(n)) = h(\hat{u}(n-1)) + H(n) [u(n) - \hat{u}(n-1)] + \dots, \quad (10)$$

式中

$$H(n) = \left. \frac{\partial h(w)}{\partial w} \right|_{w=\hat{u}(n-1)}. \quad (11)$$

采用一级近似,系统的静态方程(7)(8)可以近似为线性系统,

$$u(n) = u(n-1) = w_0, \quad (12)$$

$$d(n) = H^T(n)u(n) + \zeta(n) + \epsilon(n). \quad (13)$$

这里

$$\zeta(n) = h(\hat{u}(n-1)) - H^T(n)\hat{u}(n-1). \quad (14)$$

假设 $\zeta(n)$ 与 $u(n)$ 无关,可以得到估算 $\hat{u}(n)$ 的递归方程^[16],

$$\hat{u}(n) = \hat{u}(n-1) + K(n) [d(n) - h(\hat{u}(n-1))], \quad (15)$$

$$K(n) = \lambda^{-1} P(n-1) H(n) [I + \lambda^{-1} H^T(n) \times P(n-1) H(n)]^{-1}, \quad (16)$$

$$P(n) = \lambda^{-1} P(n-1) - \lambda^{-1} K(n) H^T(n) P(n-1). \quad (17)$$

这里 $K(n)$ 是增益矩阵, $P(n)$ 是 RLS 算法中的协方差矩阵,满足

$$P(n) = [\phi(n)]^{-1}, \quad (18)$$

$$\phi(n) = \lambda \phi(n-1) + H^T(n) H(n). \quad (19)$$

3.3. 网络结构优化——基于 RLS 的自适应删剪算法

用神经网络解决实际问题时,如何找到最佳的网络规模是一个重要的问题^[17].网络规模过大,虽然网络收敛速度很快但推广能力很差,规模太小,网络无法收敛.因此,针对一个实际问题,找到网络的最佳规模是一个必须解决的问题,解决该问题的有效方法是权重删剪算法.该算法采用较大的初始网络结构,然后通过删除不重要的权重,进而优化网络的拓扑结构,得到最佳的网络规模.结合上述 RLS 算法,我们把网络的训练和权重的删剪算法结合在一起用于网络的训练和拓扑结构的优化.

当第 n 个训练样本送入网络之后,定义网络的能量函数为

$$E = \frac{1}{2} \left[\sum_{j=1}^n (d(j) - h(w, x(j)))^2 + w^T P(0)^{-1} w \right]. \quad (20)$$

依据经典权重删剪方法^[18],权重矢量的波动 $\Delta \hat{w}$ 引起能量变化为

$$\Delta E = \Delta w^T \left. \frac{\partial E(w)}{\partial w} \right|_{w_0} + \frac{1}{2} \Delta w^T \left. \frac{\partial^2 E(w)}{\partial w^2} \right|_{w_0} \Delta w + \vartheta, \quad (21)$$

式中 w_0 是使能量函数局域最小化的解, ϑ 为高次项. 由于网络已经收敛 (21) 式等号右端第一项和高次项可以忽略不计. 假设协方差矩阵的初始量为一对角矩阵 $P(0) = \delta^{-1} I$, 其中 I 为单位矩阵, $\delta > 0$, 那么权重变化 $\Delta \hat{w}$ 引起能量的变化为

$$\Delta E = \frac{1}{2} \Delta w^T P(\infty)^{-1} \Delta w. \quad (22)$$

权重 w_j 的重要性由下式计算:

$$\Delta E_j = \frac{1}{2} [P(\infty)^{-1}]_{jj} \hat{w}(\infty)_j^2, \quad (23)$$

式中 $\hat{w}(\infty)$ 和 $P(\infty)$ 分别为网络收敛后的权重和协方差矩阵 $[P(\infty)^{-1}]_{jj}$ 表示 $P(\infty)^{-1}$ 的第 j 个对角元素.

基于以上公式,下面给出删剪过程.

(I) 用 RLS 算法训练网络,利用收敛的权重 $w(\infty)$ 协方差矩阵 $P(\infty)$ 和能量 E 根据 (23) 式对所有权重的重要性进行估算. 对所有的权重按重要性从小到大进行排列,设排列的序号为 $\{\pi_i\}$, 满足 $\Delta E_{\pi_m} \leq \Delta E_{\pi_k} (m < k)$.

$$(II) \text{ 设 } \Delta w_{\pi_k} = \begin{cases} [w]_{\pi_k} & 1 \leq k \leq k', \\ 0 & k > k', \end{cases}$$

由 (22) 式估算由于权重 w_{π_1} 到 $w_{\pi_{k'}}$ 的删剪而引起的能量变化 ΔE .

(III) 如果 $\Delta E \leq \alpha E$ (其中 $0 < \alpha < 1$), 则 $k' = k' + 1$, 回到 (II), 否则, 删除网络中 w_{π_1} 到 $w_{\pi_{k'-1}}$ 的权重.

值得指出的是,对于 3 层前馈神经网络,若只含有一个输出单元,则删去 1 个隐层到输出层的权重,就等于删去了 1 个隐层神经元,从输入层到该隐层单元的权重也将全部被删除. 在这种情况下,删剪算法起到了删除隐层神经元的作用.

通过删除隐层神经元,可使在线学习的计算复杂度降低. 对于多层前馈神经网络,假设神经元的数目为 M , 连接到每个神经元的权重数目平均为 N , 根据文献[17],RLS 算法的时间、空间复杂度都为 $O(MN)$. 对于单输出的 3 层网络,若输入维数较大,则易看出该计算复杂度可近似为 $O(HN)$, 其

中 H 为隐层单元数目. 设删剪前后隐层单元的数目分别为 $H_0, H_1 (H_0 > H_1)$, 则计算复杂度之比为 $\left(\frac{H_0}{H_1}\right)^2$.

4. 计算机模拟结果

4.1. 离线学习

在计算机模拟中,我们采用的数据是 Bell 实验室提供的电影“Star Wars”经过压缩编码的帧数据量序列,它的时间长度约为两个小时,速度为每秒 24 帧,共计 171000 帧,只采用了帧内压缩方式,每帧的压缩算法类似于 JPEG 标准. 从前 10000 个数据中随机抽取连续的 1000 个数据作为离线学习样本. TDNN 网络的初始结构是 12-8-1,即输入层 12 个神经元,隐藏层 8 个神经元,输出层 1 个神经元,采用 RLS 方法进行训练. RLS 方法的收敛速度很快,在多组实验中网络学习基本都在 10 个循环内收敛. 为了进行比较,我们用 BP 算法对上述数据进行了模拟,图 2 是 RLS 与 BP 的收敛曲线的比较. 从图 2 可以看出,RLS 比 BP 的收敛速度快得多,并且收敛之后的误差 RLS 也比 BP 小很多. 网络收敛以后,对权重按重要性排序,得到了能量变化随删剪权重数目变化的关系(图 3). 由图 3 可以看出,前 10 个不重要的权重被删除不会引起能量的明显变化,其中 3 个权重为隐藏层到输出层的权重. 由于输出层为一个神经元,删掉 3 个隐层到输出层权重就意味着 3 个隐藏层单元被删去,因此网络的规模变成 12-5-1,删去 3 个隐层单元后使在线学习阶段的计算复杂度下降到删剪前的 $\left(\frac{5}{8}\right)^2 = 0.391$. 这说明基于 RLS 的自适应

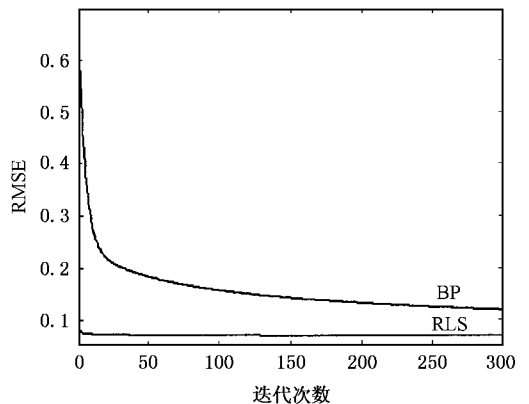


图 2 用 RLS 与 BP 算法训练 TDNN 的收敛曲线

删剪算法能有效删除不重要的权重,进而删除隐层神经元,实现网络拓扑结构的优化.

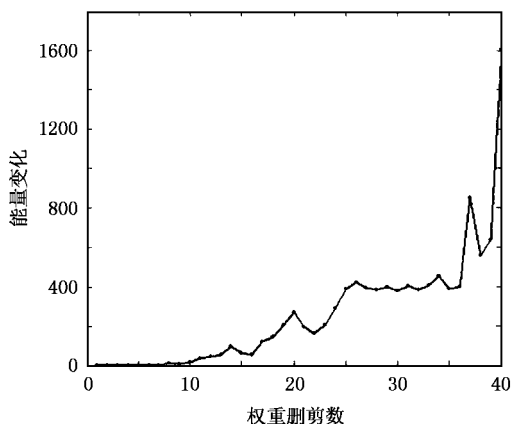


图3 能量与权重删剪数目的关系曲线

4.2. 预测过程中的在线学习

本阶段选取通信量序列中 20000—40000 这 20000 个数据作为检验样本,采用 RLS 算法进行在线预测中的再学习,遗忘因子取为 0.999.引入遗忘因子的重要意义在于,在兼顾离线学习获得的先验知识的基础上,能够遗忘部分在线学习过程中得到的 VBR 视频通信量的局部统计特征,避免网络出现学习饱和而无法根据当前通信量统计特征进行权重更新的情况.二次学习后网络的预测曲线如图 4 所示,得到的相对均方误差(relative mean square error, 简记为 RMSE)为 0.0018465,这里的 RMSE 定义为

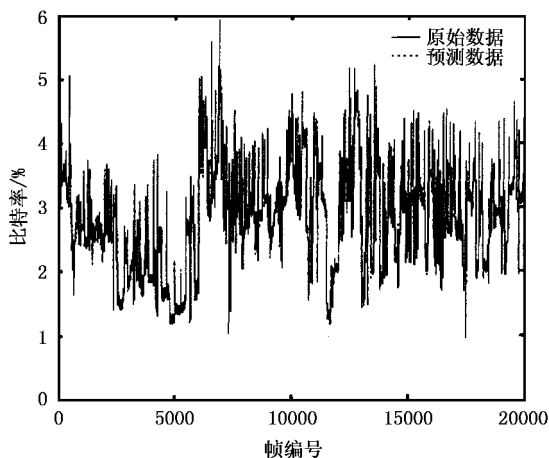


图4 网络的预测曲线

$$\text{RMSE} = \frac{\sum_n e^2(n)}{\sum_n x^2(n)}; \quad (24)$$

为了与其他方法比较,表 1 中同时列出了其他方法对同一数据的模拟结果.从表 1 可以看到,对同样的 VBR 通信量序列,我们提出的自适应 TDNN 模型得到了较高的预测精度.表 1 TDNN 模型的第二项为离线学习阶段没有经删剪处理的结构(即 12-8-1 结构)作为在线再学习的初始网络结构所得到的 RMSE,与 TDNN 模型第一项结果的比较可以看出,自适应删剪算法优化了网络拓扑结构,提高网络的推广能力.

表 1 不同方法预测的 RMSE 值比较

预测算法	RMSE 值
自适应 TDNN (12-5-1)	0.0018465
自适应 TDNN (未经删减:12-8-1)	0.0018714
PRNN	0.002900
离线 TDNN	0.0041337
线性模型	0.006900

表 1 中的 PRNN 模型每一次在线修改权重时需要进行多次迭代,平均迭代次数约为 57 次,而我们提出的自适应 TDNN 模型在线学习阶段,每一次在线修改权重只需要执行一遍(15)~(17)式固定的矩阵运算即可,因此计算复杂度大大降低.相对于表 1 所列的线性方法和静态的 TDNN 模型,我们提出的自适应 TDNN 模型的预测精度有较大提高.

5. 结 论

本文提出一种用于 VBR 视频通信量预测的自适应神经网络模型.该模型分为离线学习、在线学习两个阶段:离线学习采用基于 RLS 的自适应训练及删剪算法,通过删除隐层神经元,得到一个结构紧凑的网络作为下一步在线预测的先验知识;在线预测过程中,网络根据当前预测误差,用 RLS 算法对网络进行再学习,使网络能够在兼顾先验知识的前提下适应当前通信量的变化情况,提高预测精度.VBR 通信量预测的计算机模拟结果表明,我们提出的模型能够对 VBR 视频通信量进行实时、精确的预测.

- [1] Tekalp A M 1998 *Digital Video Processing* (Englewood Cliffs : Prentice Hall)
- [2] Adas A 1998 *IEEE/ACM Trans. Networking* **6** 635
- [3] Yoo S J 2002 *IEEE Trans. Broadcasting* **48** 10
- [4] Chang S J , Zhang W W , Shen J Y *et al* 1998 *Acta Phys. Sin.* **47** 1102 (in Chinese)[常胜江、张文伟、申金媛等 1998 物理学报 **47** 1102]
- [5] Shen J Y , Chang S J , Jia J *et al* 1998 *Acta Phys. Sin.* **47** 1966 (in Chinese)[申金媛、常胜江、贾 佳等 1998 物理学报 **47** 1966]
- [6] Moh W M , Chen M J , Chu N M *et al* 1995 *Comp. Commun.* **18** 563
- [7] Su X X , Chang S J , Ma T B *et al* 2003 *J. Optoelectronics · Laser* **14** 842 (in Chinese)[苏晓星、常胜江、马天兵等 2003 光电子·激光 **14** 842]
- [8] Chong S , Li S Q , Ghosh J 1995 *IEEE J. Selec. Areas Commun.* **13** 12
- [9] Chong P R , Hu J T 1997 *IEEE J. Selec. Areas Commun.* **15** 1087
- [10] Lin G , Liu Z M 1999 *Acta Electr. Sin.* **27** 1 [林 刚、刘泽明 1999 电子学报 **27** 1]
- [11] Doulamis A D , Doulamis N D , Kollias S D 2003 *IEEE Trans. Neural Networks* **14** 150
- [12] Maglaris B , Anastassiou D , Sen P *et al* 1988 *IEEE Trans. Commun.* **36** 834
- [13] Garrett M W , Willinger W 1994 *ACM SIGCOMM Comp. Commun. Rev.* **24** 269
- [14] Grossglauser M , Bolot J C 1999 *IEEE/ACM Trans. Networking* **7** 629
- [15] Haykin S 1989 *Adaptive Filtering Theory* (Englewood Cliffs : Prentice Hall)
- [16] Leung C S *et al* 1996 *IEE Electr. Lett.* **32** 2152
- [17] Chen S , Chang S J , Yuan J H *et al* 2001 *Acta Phys. Sin.* **50** 674 [陈 戎、常胜江、袁景和等 2001 物理学报 **50** 674]
- [18] Lecun Y , Denker J S , Solla S A 1989 *Adv. Neural Inform. Proc.* **2** 598

Adaptive training and pruning algorithm for variable bit rate video traffic prediction *

Xiong Tao Chang Sheng-Jiang[†] Shen Jin-Yuan Zhang Yan-Xin

(Institute of Modern Optics , Key Laboratory of Opto-electronic Information Science and Technology of Ministry of Education , Nankai University , Tianjin 300071 , China)

(Received 8 October 2003 , revised manuscript received 25 November 2004)

Abstract

An adaptive training and pruning algorithm for variable bit rate (VBR) video traffic prediction is proposed in this paper . By simulation and comparison , the adaptive neural network model proposed in this paper is shown to be promising and practically feasible in obtaining the best adaptive prediction of real-time VBR video traffic .

Keywords : video traffic prediction , neural network , recursive least square , pruning algorithm

PACC : 8730

* Project supported by the Key Program of the Natural Science Foundation of Tianjin (Grant No. 023800811) , the National Natural Science Foundation of China (Grant Nos. 60277022 , 60477009) , the Doctorate Foundation of the Ministry of Education of China (Grant No. 20030055022) and the Innovation and Technology Fund of Nankai University , China .

[†] Corresponding author . E-mail : sjchang@nankai.edu.cn