

用支持向量机网络实现 VBR 视频通信量的预测

李素梅, 张延, 常胜江

(南开大学信息技术科学学院, 天津 300071)

摘 要: VBR (Variable Bit Rate) 视频信号具有时变性、非线性和突发性等特点, 实现该信号通信量的高精度预测是提高信息传输速度和提高网络带宽资源利用效率的重要手段. 针对以上问题, 本文提出了一种用于 VBR 视频通信量预测的差分输入支持向量机 (SVM: Support Vector Machine) 网络模型. 该网络模型采用结构风险最小化准则, 在最小化经验风险的同时, 尽量缩小模型预测误差的上界, 从而使网络模型具有更好的推广能力. 实验结果表明: 支持向量机网络模型的预测误差为 0.0018 而梯度径向基函数 (Gradient Radial Basis Function GRBF) 神经网络模型的预测误差为 0.0029. 可以看出, 支持向量机网络模型的预测精度要比 GRBF 网络模型的预测精度高出大约 40%.

关键词: VBR 视频通信量; 支持向量机; 结构风险

中图分类号: TP183 TP18 TP391.4 **文献标识码:** A **文章编号:** 0372-2112 (2006) 02-0210-04

VBR Video Traffic Prediction Based on the SVM Networks

LI Sumei ZHANG Yan-xin CHANG Sheng-jiang

(College of Information Technical Science, Nankai University, Tianjin 300071, China)

Abstract A support vector machine neural network is proposed for performing VBR video traffic prediction. It takes differential signal as the input of the network. According to the criteria of structural risk minimization of SVM, the errors between sample-data and model-data are minimized and the upper bound of predicting error of the model is also decreased simultaneously so that the ability of generalization of the model is much improved. The simulation results show that the prediction error of SVM neural networks is 0.0018, while the prediction error of GRBF neural networks is 0.0029. In prediction precision, SVM neural networks model extended 40% than GRBF model.

Key words VBR video traffic; support vector machine (SVM); structure risk

1 引言

ATM (Asynchronous Transfer Mode), 即异步传输模式, 它能够支持包括语音、数据、图像和视频在内的各种多媒体业务. 经视频压缩产生的 VBR 视频信号具有不确定性、非线性和突发性等特点, 在网络中传输需要占用很多的带宽资源, 实现 VBR 视频通信量的精确预测是提高信息传输速度和提高网络带宽资源利用效率的重要手段. 线性预测方法^[1-2]是其中的方法之一, 由于它的计算复杂度较低, 适用于在线学习方式, 但用线性模型预测非线性的 VBR 视频信号在理论上存在不足, 预测精度较低. 神经网络作为一种新的预测方法^[3], 以其良好的非线性映射能力、灵活有效的学习方式等特点, 在预测领域的应用中表现出较大的优势和潜力^[4-5]. 较早使用的是 BP (back prop-

agation) 网络, RBF (Radial Basis Function) 网络等传统神经网络. 但是它们大多处理的是平稳的非线性时间序列, 不具备时变描述特征, 因此预测的精度是有限的. 文献 [6] 采用梯度径向基函数 (GRBF: Gradient Radial Basis Function) 网络模型实现 VBR 视频通信量的预测, 由于其输入采用时间序列的差分信号, 消除了时间序列的由于局部平均值变化而造成的不稳定性, 从而提高预测精度. 但是, 由于传统神经网络极易陷入局部极小值, 网络结构的设计过分依赖于设计者的经验和先验知识, 单纯追求经验风险最小等缺陷, 致使网络模型的推广能力减弱. 本文提出了一种用于 VBR 视频通信量预测的差分输入支持向量机网络模型. 该网络模型基于统计学习理论, 采用结构风险最小化准则, 在最小化经验风险的同时, 尽量缩小模型预测误差的上界, 从而使网络模型具有更好的推广能力. 实验结果

收稿日期: 2005-05-10 修回日期: 2005-09-06

基金项目: 天津市自然科学基金 (No. 023800811); 国家自然科学基金 (No. 60277022); 博士点基金 (No. 20030055022); 国家自然科学基金 (No. 60477009)

© 1994-2010 China Academic Journal Electronic Publishing House. All rights reserved. <http://www.cnki.net>

表明: 支持向量机网络模型的预测误差为 0.0018, 而梯度径向基函数神经网络模型的预测误差为 0.0029, 可以看出, 支持向量机网络模型的预测精度要比 GRBF 网络模型的预测精度高出大约 40%。

2 VBR 视频通信量序列的非线性自回归模型

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^[7,8]。利用前馈神经网络进行通信量预测是基于非线性自回归 (NAR: Nonlinear Autoregressive) 模型, 即:

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

其中, $\{x(n)\}$ 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 0 独立、有确定分布的随机过程。式 (1) 表明, 当前时刻流量 $\{x(n)\}$ 是过去流量 $x(n-1), x(n-1) \dots$ 的一个函数, 因此, 如果函数 $f(\cdot)$ 已知, 就可以从过去的通信量预知将来的通信量, 但困难在于式 (1) 中的非线性函数是未知的, 这一函数的映射过程可以通过支持向量机网络模型的学习得到。

3 支持向量机网络模型及算法

3.1 支持向量机网络模型简介

SVM 不仅可以求解模式分类任务^[9~11], 而且还可以用于求解非线性回归问题^[9,10]。为了构造支持向量机逼近网络的期望输出, Vapnik 最早提出了 ε -不敏感损失函数:

$$L_{\varepsilon}(d, y) = \begin{cases} |d - y| - \varepsilon & |d - y| \geq \varepsilon \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

如图 1 所示, 其 ε 是指定的参数, d 为期望输出, y 为网络输出。处理支持向量机在非线性回归问题中的思路与模式分类中支持向量机的处理方法十分相似, 首先考虑用线性回归函数 $f(x)$

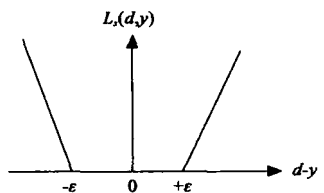


图 1 ε -不敏感损失函数

$= w \cdot x + b$ 拟合数据 $\{x_i, d_i\}, i = 1, 2, \dots, n, x_i \in R^d, d_i \in R$ 的问题, 并假设所有的训练数据可用线性函数无误差地以精度 ε 拟合, 即:

$$\begin{cases} d_i - w \cdot x_i - b \leq \varepsilon \\ w \cdot x_i + b - d_i \leq \varepsilon \end{cases}, \quad i = 1, 2, \dots, n \quad (3)$$

与最优分类面中最大化分类间隔相似, 这里控制函数集复杂性的方法是使回归函数最平坦, 它等价于最小化 $\frac{1}{2} \|w\|^2$ 。考虑到允许拟合误差的情况, 引入两组非负的松弛因子 $\xi_i \geq 0$ 和 $\bar{\xi}_i \geq 0$ 则条件式 (3) 变成:

$$\begin{cases} d_i - w \cdot x_i - b \leq \varepsilon + \xi_i \\ w \cdot x_i + b - d_i \leq \varepsilon + \bar{\xi}_i \end{cases}, \quad i = 1, 2, \dots, n \quad (4)$$

需要优化的目标函数为:

$$\min \Phi(w, \xi, \bar{\xi}) = \frac{1}{2} w^T w + C \left[\sum_{i=1}^n (\xi_i + \bar{\xi}_i) \right] \quad (5)$$

其约束条件为式 (4)。常数 $C > 0$ 控制对超出误差 ε 的惩罚程度。与模式分类相同, 通过 Lagrange 乘子 (α_i 和 α_i^*) 得到上述问题的对偶问题,

$$\begin{aligned} \max Q(\alpha_i, \alpha_i^*) = & \sum_{i=1}^n d_i (\alpha_i - \alpha_i^*) - \varepsilon \sum_{i=1}^n (\alpha_i + \alpha_i^*) \\ & - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n (\alpha_i - \alpha_i^*) (\alpha_j - \alpha_j^*) \langle x_i, x_j \rangle \end{aligned} \quad (6)$$

其约束条件为:

$$\begin{cases} \sum_{i=1}^n (\alpha_i - \alpha_i^*) = 0 \\ 0 \leq \alpha_i, \alpha_i^* \leq C, i = 1, 2, \dots, n \end{cases} \quad (7)$$

相应的回归函数为

$$f(x) = (w \cdot x) + b = \sum_{i=1}^n (\alpha_i - \alpha_i^*) \langle x_i, x \rangle + b^* \quad (8)$$

这里 α_i 和 α_i^* 只有小部分不为零, 它们对应的样本就是支持向量, 一般是在函数变化比较剧烈位置上的样本。而且这里只涉及内积运算, 与维数无关。要实现非线性函数的拟合, 只要用核函数 $K(x_i, x_j) = \langle \varphi(x_i), \varphi(x_j) \rangle$ 替代式 (6)、(8) 中的 $\langle x_i, x_j \rangle$ 即可。其网络结构如图 2 所示。

支持向量机的训练目的是求非线性回归函数式 (8), 其本质就是求解二次规划 (QP: Quadratic Programming) 问题, 相当于在约束条件式 (7) 下求式 (6) 的最大值,

将这个问题可以表示成矩阵形式

$$Q(\alpha) = \frac{1}{2} \alpha^T H \alpha + f^T \alpha \quad (9)$$

其约束条件应该为

$$\begin{cases} Aeq \cdot \alpha = 0 \\ 0 \leq \alpha \leq C \cdot 1 \end{cases} \quad (10)$$

其中 $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n, \alpha_1^*, \alpha_2^*, \dots, \alpha_n^*)^T$,

$H = \begin{bmatrix} K & -K \\ -K & K \end{bmatrix}, f = (\varepsilon - d_1, \varepsilon - d_2, \dots, \varepsilon - d_n, \varepsilon + d_1, \varepsilon + d_2, \dots, \varepsilon + d_n)^T, Aeq = (1, 1, \dots, 1, -1, -1, \dots, -1)$, (维数为 $2n$), $C \cdot 1 = (C, C, \dots, C)^T$ (维数为 $2n$), $K_{ij} = K(x_i, x_j)$ 。直接求解这个二次规划问题就可得到所求问题的支持向量机, 通过回归函数式 (8) 可求出 b^* 。

3.2 支持向量机网络的输入

在本文采用的支持向量机网络中, 输入矢量是由原始信号的差分产生, 其在 n 时刻的输入为:

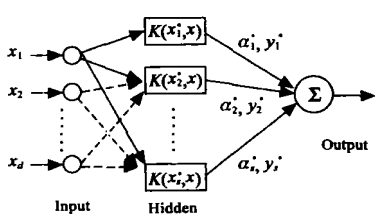


图 2 SVM 的网络模型

$$X'(n) = X(n) - X(n-1) = \begin{pmatrix} x(n-1) - x(n-2) \\ x(n-2) - x(n-3) \\ \vdots \\ x(n-M) - x(n-M-1) \end{pmatrix} \quad (11)$$

$X'(n)$ 的各分量是过去 M 个时刻信号的变化。

3.3 支持向量机网络的核函数

用于非线性回归支持向量机核函数的选取与模式分类相同,同样可以选取以下几种函数形式:(1) d 次多项式核函数 $K(x, x_i) = [(x^T x_i) + 1]^d$;

(2)径向基函数核函数 $K(x, x_i) = \exp\left[-\frac{\|x - x_i\|^2}{2\sigma^2}\right]$;

(3)3层神经网络核函数 $K(x, x_i) = \tanh[v(x^T x_i) + a]$ 。

4 实验及结果分析

在实验中,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^[12],它的时间长度约为两个小时,速度为

24帧/秒,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。图3为相

应的通信量数据曲线。

从前10000个数据

中随机抽取连续的

500个数据经归一化

后作为学习样本。根

据所采用数据的特点,

本文选取支持向量

机网络的输入维数

为12。参数 C 、 σ

和 ε 分别取6000、0

006和0.01。网络对

500个训练数据的拟

合曲线如图4所示。

参数 C 、 σ 和 ε 的选取对

曲线的拟和精度影响

较大, C 越大,支持

向量的个数越少,宽

度 σ 越大拟合精度

越差, ε 越小则拟

合精度就越好。

在测试过程中,

为了与其它模型^[6]

作比较,同样选取

20001~40000之间的

20000个数据为测

试数据。由于预测

阶段我们无法预先

知道序列的平均值

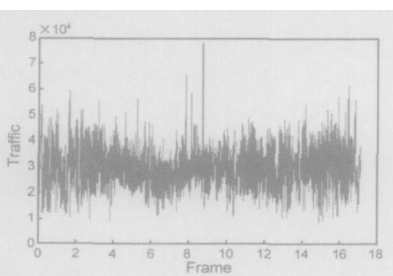


图3 Star Wars压缩视频业务量时间序列曲线

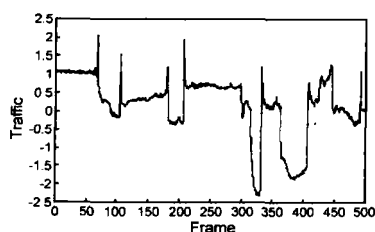


图4 500个训练数据的拟合曲线

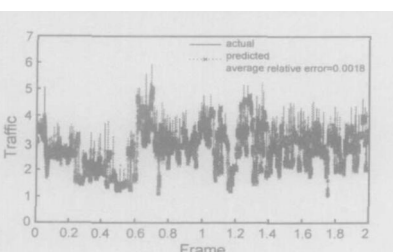


图5 SVM模型预测曲线示意图

与方差,因此,在这一阶段,采用递归迭代的方法进行在线归一化预处理,最终所得的预测曲线如图5所示。在测试过程中参数 σ 的取值也会影响测试效果,我们选取参数 σ 取0.01。

为了定量地分析网络的预测精度,与文献[6]相同,本文采用相对均方误差来表示网络的预测精度,计算公式为:

$$mse = \frac{\sum_n [e(n)]^2}{\sum_n [x(n)]^2} \quad (12)$$

采用式(12),GRBF、RBF和SVM得到的预测误差分别为0.0029、0.08575和0.0018。为了同其他几种网络模型进行比较,还对其他两种方法采用同一数据进行模拟,所得到的相对均方误差也同时列入表1中。可以看到,SVM和GRBF网络的预测精度比其它模型有了量级的提高,而SVM的预测精度最高,主要的原因是差分输入能有效地消除由于局部平均值的变化而造成的不稳定性的影响,并且SVM采用结构风险最小化原则,在最小化训练样本点误差的同时,缩小模型预测误差的上界,从而提高了模型的泛化能力。因而能给出比TDNN[13]模型、线性LMS模型、RBF更精确的预测效果。

SVM方法的不足之处在于,参数 C 、 σ 和 ε 的选取会影响网络的性能,如何选取这些参数和选取什么样的值仍是一个有待解决的问题。

表1 预测误差比较

Prediction Algorithm	Mean Relative Error
SVM	0.0018
GRBF	0.0029
Linear LMS	0.016764
Offline TDNN	0.05006
RBF	0.08575

5 结论

本文提出一种用于VBR视频通信量预测的支持向量机网络模型。该模型采用结构风险最小化准则,在最小化经验风险最小的同时,尽量缩小模型预测误差的上界,使得网络有更好的推广能力。并且网络的输入采用差分形式,消除由于局部平均值随时间变化而造成的不稳定性。实验结果表明,支持向量机网络模型对VBR视频通信量有更高的预测精度。

参考文献:

- [1] A Adas Using adaptive linear prediction to support real time VBR video under RCBR network service model [J]. IEEE/ACM Trans Networking 1998, 6(5): 635-644.
- [2] S J Yoo. Efficient traffic prediction scheme for real time VBR MPEG video transmission over high-speed net-

- works[J]. IEEE Trans on Broadcasting, 2002, 48(1): 10–18
- [3] Platt J A resource-allocating network for function interpolation[J]. Neural Computation, 1991, 2(3): 213–225
- [4] Foster W R, Collopy F, Ungar L H. Neural network forecasting of short noisy time series[J]. Computers Chem Enging, 1992, 16(4): 293–297
- [5] 苏晓星, 常胜江, 马天兵, 申金媛, 张延炘. 一种用于 ATM 业务量预测的神经网络模型[J]. 光电子·激光, 2003, 14(8): 842–845
- [6] 李素梅, 常胜江, 苏晓星, 等. 用于 VBR 视频通信量预测的梯度径向基函数网络模型[J]. 光电子·激光, 2004, 15(7): 814–817
- [7] M W Garett, W Willinger Analysis modeling and generation of self-similar VBR video traffic[J]. ACM SIGCOMM Computer Communication Review, 1994, 24(4): 269–280
- [8] M Grossglauser, J C Bolot On the relevance of long-range dependence in network traffic[J]. IEEE/ACM Trans Networking, 1999, 7(5): 629–640
- [9] Vapnik V. The Nature of Statistical Learning Theory [M]. New York: Springer Verlag, 1995, 5–13
- [10] Simon Haykin, Neural Networks: A Comprehensive Foundation (second Edition) [M]. Beijing: Tsinghua University Press, 2001, ISBN 7-302-04936-X/TP • 2778, 318–350
- [11] 李素梅, 张延炘, 等. 基于 SVM 的人眼注视与否的探知[J]. 光电子·激光, 2004, 15(10): 842–845
- [12] M Grossglauser, J C Bolot On the relevance of long-range dependence in network Traffic[J]. IEEE/ACM Trans Networking, 1999, 7(5): 629–620
- [13] A D Doulan, S N D Doulan, S D Kollias An adaptable neural network model for recursive non-linear traffic prediction and modeling of MPEG video sources[J]. IEEE Trans Neural Networks, January 2003, 14(1): 150–166

作者简介:



李素梅 女, 2004 年获南开大学工学博士学位, 现为南开大学自动化专业博士后, 主要研究方向为人工神经网络及智能信息处理。
E-mail: tjnksm@163.com.

张延炘 男, 教授, 博士生导师, 主要研究方向为人工神经网络与机器视觉; 承担国家和省部级科研项目 23 项; 发表学术论文 230 多篇; 获国家发明专利 4 项。E-mail: yxzhang@nankai.edu.cn