一种混沌海杂波背景下的微弱信号检测方法^{*}

行鸿彦 朱清清 徐伟

(南京信息工程大学气象灾害预报预警与评估协同创新中心,江苏省气象探测与信息处理重点实验室,电子与信息工程学院, 南京 210044)

(2013年12月17日收到;2014年1月6日收到修改稿)

基于复杂非线性系统的相空间重构理论,提出了一种基于遗传算法的支持向量机预测方法.利用改进的 自相关法和饱和关联维数法确定混沌信号的时间延迟和嵌入维,从而实现相空间重构.通过遗传算法优化支 持向量机中的惩罚系数和核函数参数,并结合支持向量机建立混沌序列的单步预测模型,从预测误差中检测 出淹没在混沌背景中的微弱信号(包括瞬态信号和周期信号).以Lorenz系统和加拿大McMaster大学利用 IPIX 雷达实测得到的海杂波数据作为混沌背景噪声进行仿真实验,结果表明该方法能够有效地从混沌背景 噪声中检测出微弱目标信号,所得的均方根误差为0.00049521 (信噪比为 – 89.7704 dB),这比传统支持向量 机方法的均方根误差 (0.049, 信噪比为 – 54.60 dB)降低了两个数量级.

关键词: 支持向量机, 遗传算法, 海杂波, 相空间重构 PACS: 05.45.Pq

DOI: 10.7498/aps.63.100505

1引言

混沌[1] 是指在确定性系统中出现的一种貌似 无规则的、类似随机的现象,如海杂波信号. 1990 年, Leung 等^[2] 分析了海杂波的特性, 确定其嵌入 维在5—9之间,最大Lyapunov指数为正,并认为 海杂波具有混沌特性. 然而, 在雷达对海进行观测 时,海杂波强度大,微弱目标信号会淹没在其中,海 杂波的大量尖峰还会造成严重虚警. 所以, 如何检 测海杂波背景中微弱目标信号是信号处理领域的 重要课题之一. 由于混沌噪声信号的复杂性, 人们 多采用传统的统计方法处理信号,如对数-正态分 布^[3]、韦尔分布^[4]、K分布^[5]等,但是这些方法不是 建立在混沌信号内在物理特性的基础上,且每种模 型都有特定的适用范围,不具有普适性. 20世纪90 年代中期,统计学习理论开始受到越来越广泛的重 视. 统计学习理论是小样本估计和预测学习的最佳 理论,与传统统计学相比,统计学习理论不仅考虑 了推理规则渐进性能要求,而且追求在有限样本条件下的最优结果.同时,在这一理论基础上发展了一种新的通用学习方法——支持向量机 (SVM)^[6,7].

基于 SVM 的微弱信号检测方法为强海杂波 中的目标信号检测提供了新思路, 对提高海面监 测水平具有重大意义, 受到国内外学者的高度重 视. 在国外, 1993年, Leung和Lo^[8]利用径向基函 数 (RBF) 神经网络对海杂波数据进行预测, 并且 引入了基于这个模型的海面微弱信号检测技术. 1997年, Mukherjee 等^[9]研究了 SVM在混沌时间 序列预测上的应用, 实验结果表明 SVM 比传统的 预测方法具有更佳的预测性能. 2002年, Haykin 等^[10]研究了海杂波的混沌特性以及保障海杂波 内在特性不受影响的去噪方法. 2008年, Kenshi 等^[11]利用极为有限的混沌时间序列重构动力系统 模型, 为相空间重构在信号检测领域的应用提供了 理论基础.

在国内,2002年,Zhu等^[12]设计了"最小平方"SVM,它在预测含有白噪声的混沌时间序列

†通讯作者. E-mail: xinghy@nuist.edu.cn

© 2014 中国物理学会 Chinese Physical Society

http://wulixb.iphy.ac.cn

^{*} 国家自然科学基金(批准号: 61072133)、江苏省产学研联合创新计划(批准号: BY2013007-02, BY2011112)、江苏省高等学校科研成果产业化推进计划(批准号: JHB2011-15)和江苏省"六大人才高峰"计划资助的课题.

上具有传统方法无法比拟的性能. 同年, 戴栋 等^[13]利用遗传算法(GA)对混沌系统的参数进行 了在线估计和校正,以Lorenz系统为例进行了数 值仿真,得到了很好的参数估计结果. 2004年,崔 万照等^[14]根据混沌动力系统的相空间重构理论, 建立了混沌时间序列的SVM 预测模型, 取得了比 较好的预测结果.同年,游荣义等[15]应用小波变 换和非线性动力方法研究了混沌信号在相空间中 的行为,证实了利用小波变换研究混沌信号是有效 的. 2005年, 叶美盈等^[16]设计了一种在线最小二 乘SVM (LS-SVM) 回归的混沌时间序列预测方法, 达到了混沌系统实时预测的目的. 2007年, 行鸿彦 和徐伟[17]利用神经网络有效地从混沌背景信号中 检测出微弱信号, 且检测门限较低. 2010年, 行鸿 彦和金天力^[18]提出了基于对偶约束的LS-SVM模 型,并将真实的海杂波作为该模型的实验数据,与 传统的SVM方法相比,其精度和检测效果有了显 著提高.

本文对基于 GA 的 SVM (GA-SVM) 检测混沌 背景下的微弱信号进行了研究,利用 SVM 和相空 间重构理论建立预测模型,并结合 GA 优化 SVM 的参数.以Lorenz 系统和海杂波数据作为实验信 号进行仿真实验,并与其他实验方法进行比较,研 究其检测门限和预测精度.

2 相空间重构理论和SVM算法

2.1 混沌时间序列的相空间重构理论

Packard 等^[19] 认为, 单变量时间序列可以重构出混沌系统的相空间. 在实际的非线性系统中可得到时间间隔为 Δt 的单变量时间序列 { x_1, x_2, \dots, x_n }, 其中, $x_j = x(t_j), t_j = t_0 + j\Delta t, j = 1, 2, \dots, N$.

根据文献 [20, 21] 的相空间重构理论,相空间 轨迹中的每一点都存在如下映射关系:

 $x_{j+D_{\rm E}\tau} = F([x_j, x_{j+\tau}, x_{j+2\tau}, x_{j+(D_{\rm E}-1)\tau}]).$ (1) 我们采用经典的Grassberger和Procaccia提出的GP算法^[22,23]求解嵌入维 $D_{\rm E}$,用改进的自相关法^[24]求解时间延迟 τ .

2.2 SVM 算法

相空间重构是预测混沌时间序列的第一步,国 内外已有多种预测方法,如自适应非线性滤波预测 法、神经网络法^[25]、SVM法等.为了得到较好的预测精度和泛化能力,我们采用Vapnik提出的SVM方法.SVM是基于结构风险最小化原则,以训练误差作为优化问题的约束条件,以最小化置信范围值作为优化目标.SVM的求解最后转化为二次规划问题上的求解,它的解是惟一的、也是全局最优的.

对于一个给定的训练集 $(x_i, y_i), i = 1, 2, \cdots, l,$ $x_i \in \mathbf{R}^n, y_i \in \mathbf{R},$ 回归估计函数 f(x) 为

$$f(x) = \omega \cdot \phi(x) + b, \qquad (2)$$

其中, ω 的维数为特征空间维数;b为偏差量, $b \in \mathbf{R}$; ϕ 是将训练集映射到一个高维特征空间的非线性 映射.

最优化问题为

$$H(\omega, b, \xi, \xi^*) = \min\left(\frac{1}{2}||\omega||^2 + C\sum_{i=1}^{l}(\xi_i + \xi_i^*)\right).$$
(3)

(3) 式应满足下列条件:

$$y_i - \omega \cdot x_i - b \leqslant \varepsilon + \xi_i,$$

$$\omega \cdot x_i + b - y_i \leqslant \varepsilon + \xi_i^*$$

$$\xi_i \geqslant 0, \quad \xi_i^* \geqslant 0,$$

其中, C为惩罚系数, C > 0; $\xi \pi \xi^*$ 为松弛变量; ε 为不敏感损失参数. 采用 Lagrange 乘子法求解这 个二次规划问题, 最后得到 SVM 回归模型为

$$\hat{\mathcal{Y}}(x) = \sum_{i=1}^{l} (\alpha_i - \alpha_i^*) K(x_i, x) + b, \qquad (4)$$

其中, α_i, α_i^* 为Lagrange乘子; $K(x_i, x)$ 为Mercer 核函数, $K(x_i, x) = \phi(x_i) \cdot \phi(x)$, 这里采用 RBF, $K(x_i, x) = \exp(||x_i - x||^2/\sigma^2)$. 另外, 这里涉及的 惩罚系数 C 和 RBF 参数 σ 是通过 GA 最优选择得 到的.

3 GA优化SVM

泛化能力是衡量SVM模型性能的重要指标, 它是指模型对未知样本预测的准确性,而惩罚参数 *C*和核函数参数σ对SVM的性能有很大影响.事 实上,对SVM参数寻优的最终目的就是提高其泛 化能力.我们采用GA^[26]进行参数寻优,GA不依 赖于目标函数的梯度信息,同时又具有全局搜索能 力,能够求得目标函数的全局最优解,从而避免了 局部最优解问题.

3.1 编码解码方式

GA中首先需要考虑的是编码解码问题,该问 题是关于实际问题空间到GA空间的映射关系,如 图1所示. 实际问题空间中的数据通过编码得到算 法空间中的基因,算法空间中的基因通过解码得到 实际问题空间中的数据.



图1 GA的编码解码方式

3.2 适应度函数

GA不依赖于目标函数的梯度信息,进化过程 完全以适应度函数为依据.为了使GA-SVM模型 的预测值和目标函数之间的均方根误差 (RMSE) 最小,我们将适应度函数 $G(C,\sigma)$ 定义为

$$G(C,\sigma) = \frac{1}{\frac{1}{l}\sum_{i=1}^{l} (T_i - g(x_i))^2},$$
(5)

其中, T 为真实值, $q(x_i)$ 为预测值.

3.3 遗传操作

遗传操作主要包括选择操作、交叉操作和变异 操作三种操作方式.

选择的目的是为了从当前群体中选出优良的 个体,适应性越强的个体被选择的概率越大.选择 操作的方法有很多,如比例选择算子法、局部选择 法、随机遍历选择法等. 另外, 我们通过选择适应度 值来计算选择概率.

通过交叉操作可以得到新一代个体,新个体继 承了父辈个体的特性. 利用二进制编码的交叉方法 包括常用的单点交叉法、多点交叉法和均匀交叉法.

变异操作是在群体中随机选择一个个体,对于 选中的个体以一定的概率随机地改变二进制数据

中某一位的值,即进行 $0 \rightarrow 1, 1 \rightarrow 0$ 的变异. 变异 发生的概率很低,通常取0.001-0.01.

3.4 GA-SVM 主要流程

GA-SVM流程如图2所示. 用GA优化SVM 参数的步骤如下:

步骤1 对随机产生的SVM参数进行编码, 由于待优化的参数只有C和 σ ,数量较少,故采用 二进制编码;

步骤2 计算适应度值,这里选用均方误差的 倒数表示适应度函数;

步骤3 以比例选择算子法完成选择,选择概 率为

$$P_i = \frac{G_i}{\sum\limits_{i=1}^{N} G_i},\tag{6}$$

其中, N 为个体数, G_i 为第i个个体的适应度, 选择 的目的是将适应度值大的个体遗传给下一代;

步骤4 对当代群体进行交叉操作和变异操 作,产生下一代群体;





4 仿真实验

为了验证本文提出的GA-SVM 预测方法的有 效性,我们进行了三个仿真实验.实验1和实验2 均采用 Lorenz 系统作为混沌背景噪声 c(n), 分别将 瞬态信号和周期信号叠加到c(n)上形成观测序列 x(n), 然后进行 GA-SVM 单步预测, 研究其预测精 度和检测门限.实验3采用加拿大McMaster大学

利用 IPIX 雷达实测得到的海杂波数据作为混沌背 景噪声,研究 GA-SVM 模型对淹没在海杂波下微 弱目标信号的检测精度.

实验1 以Lorenz 混沌系统为例进行仿真实 验. Lorenz 系统的迭代方程为

$$\dot{x} = d(y - x),$$

$$\dot{y} = -xz + cx - y,$$

$$\dot{z} = xy - rz,$$
(7)

其中, d = 16, b = 4, r = 45.92, 初始值为x = -1, y = 0, z = 1. Lorenz系统吸引子的图形如 图3所示.



图 3 Lorenz 系统吸引子图形

用步长为0.01的Runge-Kutta法求解方程(7), 将x分量作为实验的混沌时间序列.待系统完全进 入混沌状态,取1200个点作为训练GA-SVM模型 的样本,后续的1000个点作为预测集.采用经典的 GP 算法求得嵌入维 $D_{\rm E} = 5$.用改进的自相关法求 得时间延迟 $\tau = 1$.

在预测集的第505—554个点处加入幅值为 0.006的微弱瞬态信号s(n) (n为样本点),将s(n)叠 加到混沌背景信号c(n)中构成观测时间序列x(n), 信噪比 (SNR)达到—89.7704 dB, 然后进行相空 间重构和GA-SVM单步预测,预测结果的RMSE 为0.00049521. 采用 GA 得到的最优参数为C =558.6138, $\sigma = 0.6545$, 检测结果如图 4 所示. 从 图 4 (b) 可以看出, 预测集的 n = 505-554 处单步 预测误差比较大, 另外预测集的 n = 555-558 处单 步预测误差也比较大, 这是因为 $x(554) \neq c(554)$, 用 n=554 进行单步预测, 将引起预测值 $\hat{x}(558)$ 存在 较大的误差. 与文献 [18] 中的实验结果相比, GA-SVM 对瞬态信号的检测能力相对于对偶约束 LS-SVM 对瞬态信号的检测能力有了很大的加强, 包 括误差的减小和检测门限的降低. 表 1 列出了在实



图 4 GA-SVM 对瞬态信号的检测结果 (a) 瞬态信号 s(n); (b) 含瞬态信号的真实值和预测值; (c) 单步预测误差

混沌时间序列预测模型性能的对比 表1 GA-SVM 对偶约束 LS-SVM LS-SVM SVM RBF 神经网络 SNR/dB -89.7704-77.33-62.82-54.60-30.200.022 0.049 RMSE 0.0004952 0.0080 0.058

验1的情形中,不同混沌时间序列预测模型性能的 对比.从表1可以看出,本文预测方法的性能相对 于其他方法的性能有明显的提高.

实验2 设目标信号为 $s(n) = 0.00026 \sin$

 $(2\pi\nu n)$, 归一化频率为 $\nu = 0.04$, 观测信号为 x(n) = s(n) + c(n), SNR为-89.9273 dB. 按照实验1的步骤进行检测, 得到单步预测误差序列e(n), e(n)主要包含GA-SVM模型本身的预测误差和 周期信号s(n)两部分.因此,将检测淹没在强混 沌背景噪声中的周期信号问题转化为检测淹没 在单步预测误差e(n)中的周期信号问题.将单步 预测误差e(n)进行快速傅里叶变换,可以看出预 测误差的频谱在0.04处出现峰值,结果如图5 所 示.实验2预测结果的RMSE为0.00020111,采用 GA 得到的最优参数为C = 98298, $\sigma = 0.4299$, 与文献[27]中SVR的-89.65 dB检测门限和0.022 的RMSE 以及神经网络的-62.82 dB检测门限和 0.1824的RMSE 相比均有很大的提高.



图 5 GA-SVM 对周期信号的检测结果 (a) 含周期信号的真实值和预测值; (b) 单步预测误差频谱

实验3 我们采用的海杂波数据是加拿大McMaster大学的IPIX 雷达海杂波数据,该雷达发射频率为9.39 GHz,天线高度为30 m,采用VV极化方式,每组数据包含131072个采样点,天线增益为45.7 dB.

先选用不含小目标的海杂波数据,将1000个 点作为训练样本,后续的800个点作为预测样本, 实验结果如图6所示,其均方误差为0.00120698,采 用GA得到的最优参数为 $C = 76043, \sigma = 0.1272.$ 在文献[28]中,LS-SVM的均方误差为0.0137,RBF 的均方误差为0.0153,因此本实验的预测精度有了 很大的提高.

采集含目标的海杂波数据,同样选用1000 个点作为训练样本,800个点作为预测样本, 其RMSE为0.0149,采用GA得到的最优参数为 $C = 6687900, \sigma = 5.3224.$ 类似实验2的处理方 法,预测误差进行频谱分析,实验结果如图7所示,可以看到在归一化频率为0.015处出现了明显的峰值,说明混沌背景噪声中含有微弱周期信号.



图 6 GA-SVM 对不含目标信号的 IPIX 雷达海杂波的 检测结果 (a) 不含目标信号的海杂波真实值和预测值; (b) 不含目标信号的海杂波单步预测误差



图 7 GA-SVM 检测含目标信号的 IPIX 雷达海杂波的 单步预测误差频谱

实验结果表明, GA-SVM 预测方法充分利用 了 GA 的全局搜索特性和 SVM 的优点, 能够有效 预测含有不同微弱信号的混沌时间序列, 得到单 步的预测模型的预测精度明显优于传统的 SVM 神 经网络算法的预测精度.实验中先采用 GA 对海杂 波数据进行参数寻优,选择适应度大的参数, 经过 选择操作、交叉操作、变异操作后得到最优结果, 最 后进行 SVM 训练和预测. GA-SVM 方法克服了传 统参数寻优方法的计算量大、易陷入局部最优等缺 点, 较之于传统混沌背景下微弱信号检测方法而 言, GA-SVM 方法降低了检测门限和预测误差.

5 结 论

基于混沌时间序列的相空间重构和局部可预 测的特点,以及GA的参数寻优特性,我们提出了 GA-SVM 预测模型,利用该模型预测混沌背景噪 声中的瞬态信号和周期信号.在SVM模型中,惩 罚参数C和核函数参数 σ 对SVM的性能有很大影 响. 若C取值小,则对样本的惩罚小,使训练误差 变大,从而系统的泛化能力变差.若C取值大,则 会导致过学习现象. 当核函数参数 σ 的值过大时, 易出现欠学习现象, 而 σ 的值过小又会出现过学习 现象. 合理地选择参数值, 对提高SVM 的性能至 关重要. 基于梯度的优化方法只能通过试验法选取 恰当的参数值,这种方法计算量大,寻优精度不高. 我们采用GA优化参数,可以克服梯度下降法的不 足. 实验结果表明, GA-SVM 预测方法可以很好地 降低检测门限和预测误差. 下一步的主要工作是研 究有效的去噪方法以减少混沌时间序列中的高斯 白噪声,如果能在很大程度上抑制白噪声的影响, 则模型可达到更好的预测精度.

参考文献

- Zhang X D, Wang Z, Zhao P D 2008 Chin. Phys. Lett. 25 397
- [2] Leung H, Haykin S 1990 Appl. Phys. Lett. 56 593
- [3] Szajnowski W J 1976 Electron. Lett. 12 497
- [4] Anastassopoulos V, Lampropulos G 1995 IEEE Trans. AES 31 52
- [5] Conte E, Lops M, Ricci G 1994 $I\!E\!E$ Proc. F 141 116
- [6] Zhang X G 2000 Acta Autom. Sin. 26 32 (in Chinese)
 [张学工 2000 自动化学报 26 32]
- [7] Cui W Z, Zhu C C, Bao W X, Liu J H 2005 Chin. Phys. 14 922
- [8] Leung H, Lo T 1993 IEEE J. Oceanic Eng. 18 287
- [9] Mukherjee S, Osuna E, Girosi F 1997 Workshop on Neural Networks for Signal Processing VII (Piscataway: IEEE Press) p511

- [10] Haykin S, Bakker R, Currie B W 2002 Proc. IEEE 90 860
- [11] Kenshi S, Yuko N, Shinichi A 2008 Chaos Solitons Fract.38 1274
- [12] Zhu J Y, Ren B, Zhang H X, Deng Z T 2002 Proceedings of the First International Conference on Machine Learning and Cybernetics Beijing, China, November 4, 5, 2002 p364
- [13] Dai D, Ma X K, Li F C, You Y 2002 Acta Phys. Sin. 51
 2459 (in Chinese) [戴栋, 马西奎, 李富才, 尤勇 2002 物理 学报 51 2459]
- [14] Cui W Z, Zhu C C, Bao W X, Liu J H 2004 Acta Phys.
 Sin. 53 3303 (in Chinese) [崔万照, 朱长纯, 保文星, 刘君 华 2004 物理学报 53 3303]
- [15] You R Y, Chen Z, Xu S C, Wu B X 2004 Acta Phys. Sin. 53 2882 (in Chinese) [游荣义, 陈忠, 徐慎初, 吴伯僖 2004 物理学报 53 2882]
- [16] Ye M Y, Wang X D, Zhang H R 2005 Acta Phys. Sin.
 54 2568 (in Chinese) [叶美盈, 汪晓东, 张浩然 2005 物理 学报 54 2568]
- [17] Xing H Y, Xu W 2007 Acta Phys. Sin. 56 3773 (in Chinese) [行鸿彦, 徐伟 2007 物理学报 56 3773]
- [18] Xing H Y, Jin T L 2010 Acta Phys. Sin. 59 140 (in Chinese) [行鸿彦, 金天力 2010 物理学报 59 140]
- [19] Packard N H, Cratchfield J P, Farmer J D, Shaw R S 1980 Phys. Rev. Lett. 45 712
- [20] Xing H Y, Cheng Y Y, Xu W 2012 Acta Phys. Sin. 61 100506 (in Chinese) [行鸿彦, 程艳艳, 徐伟 2012 物理学报 61 100506]
- [21] You R Y, Huang X J 2011 Chin. Phys. B 20 020505
- [22] Grassberger P, Procaccia I 1983 Phys. Rev. Lett. 50 346
- [23] Lo T, Leung H 1993 $I\!E\!E$ Proc. F 140 243
- [24] Aguirre L A 1995 Phys. Lett. A 203 88
- [25] Ma Q L, Zheng Q L, Peng H, Zhong T W, Qin J W 2008 Chin. Phys. B 17 536
- [26] Wang Y N, Tan W 2003 Acta Phys. Sin. 52 2723 (in Chinese) [王耀南, 谭文 2003 物理学报 52 2723]
- [27] Du J Y, Hou Y B 2007 J. Sci. Instrum. 28 555 (in Chinese) [杜京义, 侯媛彬 2007 仪器仪表学报 28 555]
- [28] Wang F Y, Yuan G N, Xie Y J, Qiao X W 2009 Radar Sci. Technol. 7 53 (in Chinese) [王福友, 袁赣南, 谢燕军, 乔相伟 2009 雷达科学与技术 7 53]

A method of weak target detection based on the sea clutter^{*}

Xing Hong-Yan[†] Zhu Qing-Qing Xu Wei

(Collaborative Innovation Center on Forecast and Evaluation of Meteorological Disasters, Key Laboratory of Meteorological Observation and Information Processing of Jiangsu Province, College of Electronic and Information Engineering, Nanjing University of Information Science and Technology, Nanjing 210044, China)

(Received 17 December 2013; revised manuscript received 6 January 2014)

Abstract

According to the phase space reconstruction theory of nonlinear system, we propose a prediction method of support vector machine based on genetic algorithm. Using the improved autocorrelation method and Grassberger-Procaccia algorithm to determine the time delay and embedding dimension of chaotic signal, the phase space reconstruction is implemented. The penalty coefficient and the kernel function parameter of support vector machine are optimized by genetic algorithm. Combined with support vector machine, single-step prediction model of the chaotic sequence is set up, so we can detect the weak signal in chaos from the prediction error (including the transient signal and periodic signal). Lorenz attractor and the data from the McMaster IPIX radar sea clutter database are used in the simulation. The proposed method can effectively detect the weak target from chaotic signal. When the signal-to-noise ratio is -89.7704 dB in the chaotic noise background, by using the new method the root mean square error can be reduced by two orders of magnitude, reaching 0.00049521, while the conventional support vector machine can reach only 0.049 under the condition of -54.60 dB.

Keywords: support vector machine, genetic algorithm, sea clutter, phase space reconstructionPACS: 05.45.PqDOI: 10.7498/aps.63.100505

^{*} Project supported by the National Natural Science Foundation of China (Grant No. 61072133), the Production, Learning and Research Joint Innovation Program of Jiangsu Province, China (Grant Nos. BY2013007-02, BY2011112), the Industrialization of Research Findings Promotion Program of Institution of Higher Education of Jiangsu Province, China (Grant No. JHB2011-15), and the "Summit of the Six Top Talents" Program of Jiangsu Province, China.

[†] Corresponding author. E-mail: xinghy@nuist.edu.cn