文章编号:1000-582X(2007)01-0114-06

支持向量机在地震预测中的应用*

肖汉光[1,2],蔡从中[1,2],袁前飞[1,2],刘兴华[1]
(1. 重庆大学 数理学院,重庆 400030;2. 新加坡国立大学 计算科学系,新加坡 117543)

摘 要:利用支持向量机分析了发生在美国加州中部的 2 次 6 级以上地震的震前大地脉动. 通过对离地震最近的 3 个地震台站的地震数据进行震前大地脉动分析,结果表明:支持向量机能有效地区分震前大地异常脉动和平静时期的大地脉动,并且随着地震的临近预报准确率逐渐增加;2 次地震的震前大地异常脉动分别始于地震前 48 h 和 12 h. 分析了加州 CI 地震台网内的 14 个地震台站记录的 2003 年 12 月 22 日发生在加州中部的 6.4 级地震所观测的震前脉动数据,发现处在震中附近的 12 个地震台站均观测到震前大地的异常脉动,且距离震中附近的断层越近,监测到震前脉动异常的几率越大. 对 3 个观测站进行连续监测,结果表明:监测到大地震($M \geq 5$)所引发的震前脉动异常的概率大于小地震($M < 5$). 因此,该方法有望发展成为地震预报的一种有效手段.

关键词:支持向量机;地震预测;震前地震波;特征提取
中图分类号:P315;TP18 **文献标识码**:A

2004 年 12 月 26 日,印尼苏门答腊岛北部发生了 9.2 级地震. 由该海底地震形成的此次印度洋大海啸席卷了印度尼西亚、斯里兰卡、印度、泰国、马尔代夫、马来西亚、缅甸、孟加拉、索马里、坦桑尼亚和肯尼亚等近 20 个国家,死亡人数达 25 万人,200 多万人无家可归,给东南亚、南亚和非洲东海岸的人民带来了巨大灾难. 这是一次有史以来最为惨烈的自然灾害之一. 据报道,全球科学家都未能预测到这次海底地震的发生. 由于地震预测的复杂性,从 19 世纪 70 年代至今地震预测和预报一直处于探索阶段,仍未形成成熟的地震预测和预报理论[1]. 目前,人们主要通过对震前先兆的观测进行地震预测和预报,如地磁及地电流观测、次声波异常观测、地壳变动的连续观测、地下水水位和化学元素含量的观测、电离层扰动观测、卫星红外异常观测、动物异常行为观测等. 例如:地震前震中附近地磁场极化强度(Z/G)增强[2],地磁扰动异常[3];地电势在震前几个星期有异常扰动[4];大震前次声波的异常;通过大震前的小震(即:前震)来预测地震的发生[5];地下水水位的涨落和所含元素含量的变化与地震有一定的关系,如地震前几天地下水水位下降和氡离子含量增高等[6];电离层中电子密度的变化与地震有着密切的关系,地震规模 5.0 级以上 5 d 前发生电离层 foF2 异常减小的概率为 74.1%,且地震的规模愈大出现概率愈大[7];地表温度(水温,地温,空气温度等)异常与地震也存在一定关系;强祖基[8],Tramutoli[9]和 Tronin[10]等分别在地震前监测到卫星红外异常现象. 目前的地震预测方法尚不能同时准确预测出地震震发的时间、位置和震级.

为了探究震前大地脉动异常的存在和研究其与发生地震之间的关系,笔者以地震仪所测大地脉动数据作为研究对象,首次提出并利用支持向量机[11]的分类原理对大地脉动异常与地震震发时间、位置和震级之间的关系进行了初步的研究.

1 支持向量机

支持向量机(Support Vector Machine, SVM)是由 Vapnik[11]及其合作者基于结构风险最小化理论提出的一种有监督的机器学习方法,被公认为小样本情况

* 收稿日期:2006-09-04
基金项目:重庆市自然科学基金资助项目(CSTC,2006BB5240);重庆大学与新加坡国立大学国际合作研究项目(R-151-000-038-592)
作者简介:肖汉光(1980-),男,重庆大学硕士研究生,新加坡国立大学访问学者、主要从事机器学习、模式识别与地震预测等研究. 蔡从中,男,研究员,电话(Tel.):86-23-65102521;E-mail:caiczh@gmail.com.

下统计及学习的经典. 由于其不需要确定各类的类条件概率密度和先验概率就能找到全局最优解，并且具有较好的泛化能力，所以被广泛地应用于诸多领域，如文本分类，手写体数字识别，语音识别，图象识别与目标探测，人脸识别，商业时序预报，水文预报，空气质量预报，地球空间物理和实验高能物理数据分析与处理，肿瘤及癌症诊断，基因微阵列表达数据分析，药物设计，蛋白质－蛋白质相互作用预测以及蛋白质结构与功能预测等[12-15].

以两类（正样本和负样本）分类问题为例，在线性可分的情况下，SVM 构建一个超平面 H：

$$\boldsymbol{W}\cdot\boldsymbol{P}+b=0, \tag{1}$$

式中，$\boldsymbol{W}$ 为权重向量，$\boldsymbol{P}$ 为特征向量；b 为一参数. 该超平面以最大边界的形式将正负样本区分开. 该超平面的构建是通过寻找向量 $\boldsymbol{W}$ 和参数 b，使其在满足条件

$$\boldsymbol{W}\cdot\boldsymbol{P}_i+b\geqslant 0\ ,(\text{对正样本},y=+1) \tag{2}$$

$$\boldsymbol{W}\cdot\boldsymbol{P}_i+b\leqslant 0\ ,(\text{对负样本},y=-1) \tag{3}$$

时，$\|\boldsymbol{W}\|^2$ 达到最小. 式中 $\boldsymbol{P}_i$ 代表第 i 个训练样本的特征向量；$\|\boldsymbol{W}\|^2$ 代表权重向量 $\boldsymbol{W}$ 的欧几里德范数；y 为样本类别标记. 在求出 $\boldsymbol{W}$ 和 b 后，通过决策函数

$$\operatorname{sign}[\boldsymbol{W}\cdot\boldsymbol{P}_j+b] \tag{4}$$

判断向量 $\boldsymbol{P}_j$ 所对应测试样本的类别. 若决策函数值为 +1，该样本属于正样本；否则，属于负样本.

在线性不可分的情况下，SVM 利用核函数 $K(\boldsymbol{P}_i,\boldsymbol{P}_j)$ 将特征向量映射到一个高维空间. 在此高维空间中，线性不可分问题被转化为线性可分问题，其决策函数为：

$$\operatorname{sign}\Big[\sum_{i=1}^{l}\alpha_i^0 y_i K(\boldsymbol{P}_j,\boldsymbol{P}_i)+b\Big], \tag{5}$$

式中，l 为训练样本数，系数 α_i^0 和 b 应使拉格朗日表达式：

$$\sum_{i=1}^{l}\alpha_i-\frac{1}{2}\sum_{i=1}^{l}\sum_{j=1}^{l}\alpha_i\alpha_j y_i y_j K(\boldsymbol{P}_i,\boldsymbol{P}_j) \tag{6}$$

达到最大值，且应满足：

$$a_i\geqslant 0\quad \text{和}\sum_{i=1}^{l}\alpha_i y_i=0. \tag{7}$$

核函数 $K(\boldsymbol{P}_i,\boldsymbol{P}_j)$ 一般取径向基函数：

$$K(\boldsymbol{P}_i,\boldsymbol{P}_j)=\mathrm{e}^{-\|\boldsymbol{P}_i-\boldsymbol{P}_j\|^2/2\sigma^2}. \tag{8}$$

2　大地脉动的特征提取

对于大地脉动时序信号，最常用的特征提取方法是短时傅里叶变换（STFT）. STFT 首先根据采样频率选择长度为 L（一般取 2^n 个采样点，n 为正整数）的窗口. 利用此窗口部分重叠地连续截取时间序列，得到 N 个短时时间序列 $\{\boldsymbol{X}_1,\boldsymbol{X}_2,\cdots,\boldsymbol{X}_N\}$（其中，$\boldsymbol{X}=(x_{j1},x_{j2},\ldots,x_{jL})$，$(j=1,\cdots,N)$. 对各个短时时间序列进行中心化处理，得到 $\{\Delta\boldsymbol{X}_1,\Delta\boldsymbol{X}_2,\cdots,\Delta\boldsymbol{X}_N\}$（其中，$\Delta\boldsymbol{X}_j=(\Delta x_{j1},\Delta x_{j2},\cdots,\Delta x_{jL})$，$(j=1,\cdots,N)$，$\Delta x_{ji}=x_{ji}-\bar{x}_i$，$(i=1,\ldots,L)$）.

为了避免采用 FFT 时产生吉布斯效应，必须对 $\{\Delta\boldsymbol{X}_1,\Delta\boldsymbol{X}_2,\cdots,\Delta\boldsymbol{X}_N\}$ 进行平滑过滤. 平滑过滤器一般选择汉明窗口（Hamming window）. 其表达式为：

$$f_i=0.54-0.46\cos\Big(2\pi\frac{i-1}{L}\Big),\ i=1,\cdots,L. \tag{9}$$

$$\Delta x_{ji}=\Delta x_{ji}f_i,j=1,2,\cdots,N,i=1,\cdots,L. \tag{10}$$

将过滤后的 $\{\Delta\boldsymbol{X}_1,\Delta\boldsymbol{X}_2,\cdots,\Delta\boldsymbol{X}_N\}$ 分别进行快速傅里叶变换后进行归一化处理，从而得到 N 个代表能谱特征的向量 $\{\boldsymbol{P}_1,\boldsymbol{P}_2,\cdots,\boldsymbol{P}_N\}$，每个向量包含 L 个 FFT 系数，其中，$\boldsymbol{P}_j=(p_{j1},p_{j2},\ldots,p_{jL})$ $(j=1,\cdots,N)$.

CI 地震台网内的宽带地震仪的采样频率一般为 20 Hz 或 40 Hz，为了统一采样频率，对采样率为 40 Hz 的站点进行 down-sampling 处理，使其采样频率为 20 Hz. L 在该文中取为 2 048. 由于得到的能谱特征向量的维数太高，除了要求内存容量大之外，还降低了计算速度. 通过观察能谱特征向量，发现其高频部分的 FFT 系数普遍较小，因而截取能谱特征向量的前 256 维作为大地脉动特征向量 $\boldsymbol{P}_j=(p_{j1},p_{j2},\cdots,p_{j256})$ $(j=1,\cdots,N)$.

3　大地脉动异常分析结果及讨论

实验数据取自发生于美国加利福尼亚中部的 2 次大地震. 一次是 2003 年 12 月 22 日 19:15:56 在 (35.7°N，－121.1°E) 处发生的 6.4 级地震（简称：震 1），另一次则是 2004 年 9 月 28 日 17:15:24 在 (35.8°N，－120.3°E) 处发生的 6 级地震（简称：震 2）. 图 1 为地震发生地点和地震台站的分布示意图，图中 2 个大圆点代表地震发生地点，19 个小圆点代表地震台网 CI 的地震台站，2 个三角形（PKD，PKD2）代表地震台网 BK 的 2 地震台站，穿过 PKD、PKD2 和 MPI 的曲线 AB 代表断层. 从图 1 可以看出，2 次地震都发生在该地区的断层 AB 附近. 各地震台站均安装了 3 个宽带地震仪（即 BHE，BHN，BHZ，分别指向东、北和地心方向）用于连续记录大地脉动. 大地脉动数据均下载于南加利福尼亚地震数据中心 SCEDC，网址为：http://www.data.scec.org/.

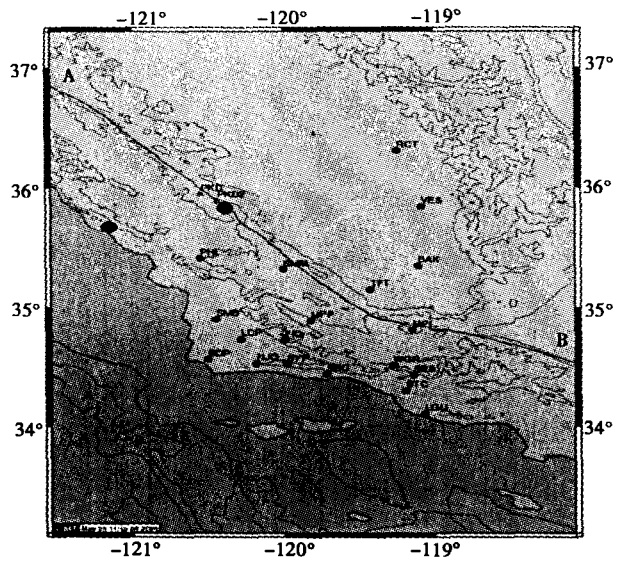


图 1 地震发生地点及地震台站分布示意图

3.1 大地脉动异常的提取

地震预测首先要解决的问题是如何准确地提取和鉴别代表地震前兆的信息. 地震震前大地脉动异常的提取是指将代表地震前兆的脉动和非震前脉动(即不能代表地震前兆的脉动或地震平静时期的脉动)区分开. 为了达到这一目的,笔者假定地震前一段时间内的大地脉动中存在震前异常脉动(即:正样本),地震平静时期(地震发生前几天或前几个月)的大地脉动为非异常脉动(即:负样本). 震 1 的正样本设为震前48 h 内的大地脉动. 由于在震 2 发生的前 22 h(2004 年 9 月 27 日 19:27:29)该区域附近(39.3° N, -133.3°E)发生过 2.7 级小震,为了研究该小震对震前脉动的影响,将震 2 的正样本设为震前 24 h 的大地脉动数据. 通过搜索 2003 和 2004 年此地区地震的发生情况,选择 6 个地震活动较为宁静时期采集负样本. 震 1 的负样本为 2003 年 9 月 8 日、2003 年 10 月 9 日和 2003 年 10 月 28 日的大地脉动数据,震 2 的负样本设为2004 年9 月 25 日的大地脉动数据. 所有数据均为离地震震中最近的 3 个地震台站(BK. PKD、CI. SMM 和 CI. PHL)的 9 个宽带地震仪所测的大地脉动数据. 正样本经过特征提取后,按时间顺序将各站点的特征向量分为 24 组(震 1)和 12 组(震 2),各组包含 2 h 3 个地震仪(BHE、BHN 和 BHZ)记录的大地脉动. 每组数据随机地分成训练集和测试集,负样本经过特征提取后得到震 1 负样本集和震 2 负样本集. 2 组负样本集分别随机地分成训练集和测试集,训练集和测试集中的负样本数相等. 分别利用支持向量机和各组训练集进行训练,然后利用对应的测试集进行分类准确率测试. 各组负样本测试分类准确率均高于 90%,震 1 和震 2 正样本测试分类准确率如图 2 所示,图中'+'、'x'和'*'号分别代表由地震台站 BK. PKD、CI. SMM 和CI. PHL采集数据的实验结果,'○'则代表合并 3 站采集数据的实验结果.

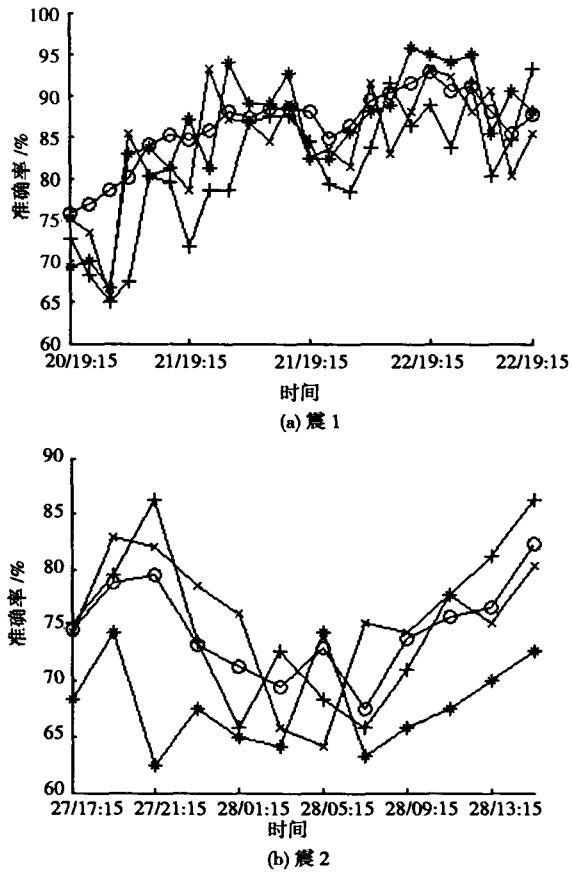
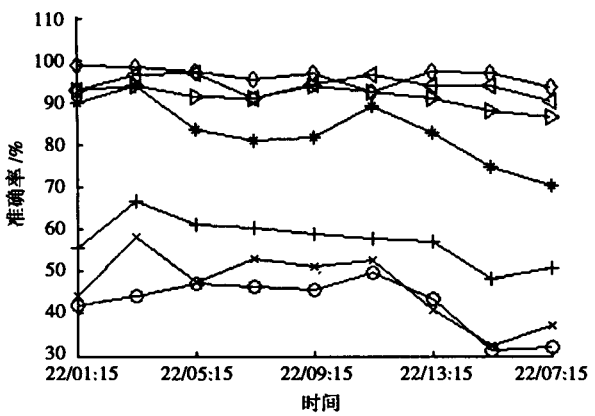


图 2 SVM 对震 1 和震 2 正样本测试集分类的准确率随时间的变化

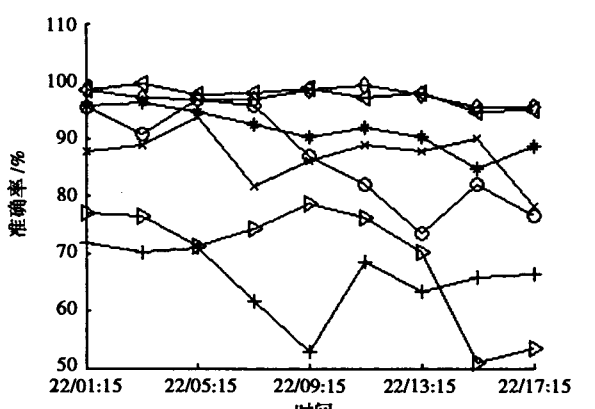
从图 2 可以看出:支持向量机能成功地将正负样本区分开来,且随着地震的临近准确率逐步增加. 由于震 2 前 22 h 有小震发生,从图 2(b)可以看出,震 2 的正样本测试集分类准确率受到余震的影响,发生于 19:15 的小震提高了分类准确率,该小震过后 8 h 内其影响逐渐减小. 在临近震 2 发生的前 12 h 内,预测的地震发生的准确率与震 1 的规律一样,随着地震的临近准确率逐渐增加. 从图 2 还可看出,3 个站点预测的趋势基本一致,3 站数据合并得到的预测结果较 3 站独立分析结果更趋稳定. 如果将分类准确率大于 70% 的正样本当作震前大地异常脉动,则震 1 的震前大地异常脉动始于震前 48 h,震 2 的震前大地异常脉动约始于震前 12 h.

3.2　震前监测模型的建立和区域震前行为分析

震 1 临震前 24 h 和震 2 临震前 12 h 的大地脉动作为新的正样本集，震 1 和震 2 的原负样本合并为新的负样本集. 将新的正负样本集合并为新的训练集. 利用支持向量机和新的训练集训练出地震震前监测模型. 该模型利用最优超平面将特征向量空间分为正样本区和负样本区，通过比较样本与超平面的相对位置，便可判断此样本属于正样本还是负样本. 在 CI 地震台网的 19 个地震台站中，除 PHL 和 SMM 2 站的数据用于训练，SMB、SES 和 STC 3 站无相应数据外，余下 14 个地震台站所测的震 1 临震前 18 h 的大地脉动数据经过特征提取后组成 14 个测试集（每个测试集对应一个站点），每个测试集按时间顺序分为 9 个测试子集，每个测试子集包含连续 2 h 内 3 个地震仪（BHE、BHN 和 BHZ）的大地脉动信息. 假定 14 × 9 个测试子集中的样本都为正样本，其 SVM 测试分类准确率如图 3 所示.



(a) 地震台站 RCT、VES、BAK、LCP、TFT、MPP 和 FIG



(b) 地震台站 SDP、NJQ、SYP、SBC、LGU、WGR 和 MPI

图 3　SVM 对各个地震台站的震 1 正样本测试集分类准确率随时间的变化

图 3（a）是地震台站 RCT、VES、BAK、LCP、TFT、MPP 和 FIG 的正样本测试集分类准确率随时间的变化情况，图中‘ + ’、‘ x ’、‘ * ’和‘ ○ ’分别代表站点 RCT、VES、BAK 和 LCP，右三角形、菱形和左三角形代表站点 TFT、MPP 和 FIG. 图 3（b）是地震台站 SDP、NJQ、SYP、SBC、LGU、WGR 和 MPI 的正样本测试集分类准确率随时间的变化情况，图中‘ + ’、‘ x ’、‘ * ’和‘ ○ ’分别代表站点 SDP、NJQ、SYP 和 SBC，右三角形、菱形和左三角形代表站点 LGU、WGR 和 MPI.

从图 3 可以看出：离断层带较远的站点 RCT、VES、LCP、SDP 和 LGU 的测试分类准确率较低，并且准确率随时间变化幅度较大；离断层带较近的站点 BAK、NJQ、SYP 和 SBC 的测试分类准确率比远站点的高，且稳定性好. 距离断层带最近的站点 TFT、MPP、FIG、MPI 和 WGR 的准确率与其他站点相比最高，且稳定性最好.

3.3　利用震前监测模型进行震区连续监测

随机选择 FIG、TFT 和 BAK 3 站点对震 2 震前 12 d 的大地脉动进行连续监测，以研究 SVM 震前监测模型对不同震级和不同距离的地震的监测效果. 各站点震前 12 d 的脉动数据经过特征提取后，按地震台站（FIG、TFT 和 BAK）、地震仪（BHE、BHN 和 BHZ）和时间顺序分为 3 × 3 × 12 × 12 个测试子集，每个测试子集包含 2 h 内 1 个地震台站的 1 个地震仪（BHE 或 BHN 或 BHZ）记录的大地脉动. 假定 3 个站点的 3 × 3 × 12 × 12 组数据都为正样本. 图 4（a）为 2004 年 9 月 16 日 17:15:00 至 2004 年 9 月 28 日 17:15:00 发生在纬度范围为（30°N ~ 40°N）和经度范围为（ - 130°E ~ - 110°E）区域的地震震级图. 3 个站点 9 个地震仪的测试分类准确率如图 4（b）（c）（d）所示，图中五角星、星号和圆圈分别代表地震仪 BHE、BHN 和 BHZ 的测试集分类结果.

从图 4 中可以看出：震前大地脉动异常出现在 6 个时间段（17 日 17 时 - 18 日 17 时、19 日 17 时 - 20 日 17 时、21 日 17 时 - 22 日 02 时、23 日 17 时 - 24 日 17 时、26 日 02 时 - 27 日 02 时和 27 日 18 时 - 28 日 17 时）. 在这 6 个时间段内，开始存在一个较短时间的异常平静期，随后 3 个地震台站的测试分类准确率增大. 处在断层上的地震台站 FIG 和 TFT 的观测结果基本相同. 离断层较远的地震台站 BAK 在大于 5 级的地震前测试分类准确率达 80% 以上. 对于震级小于 5 级的地震，其测试分类准确率均低于 80%（除 26 日 17 时 - 27 日 17 时外）.

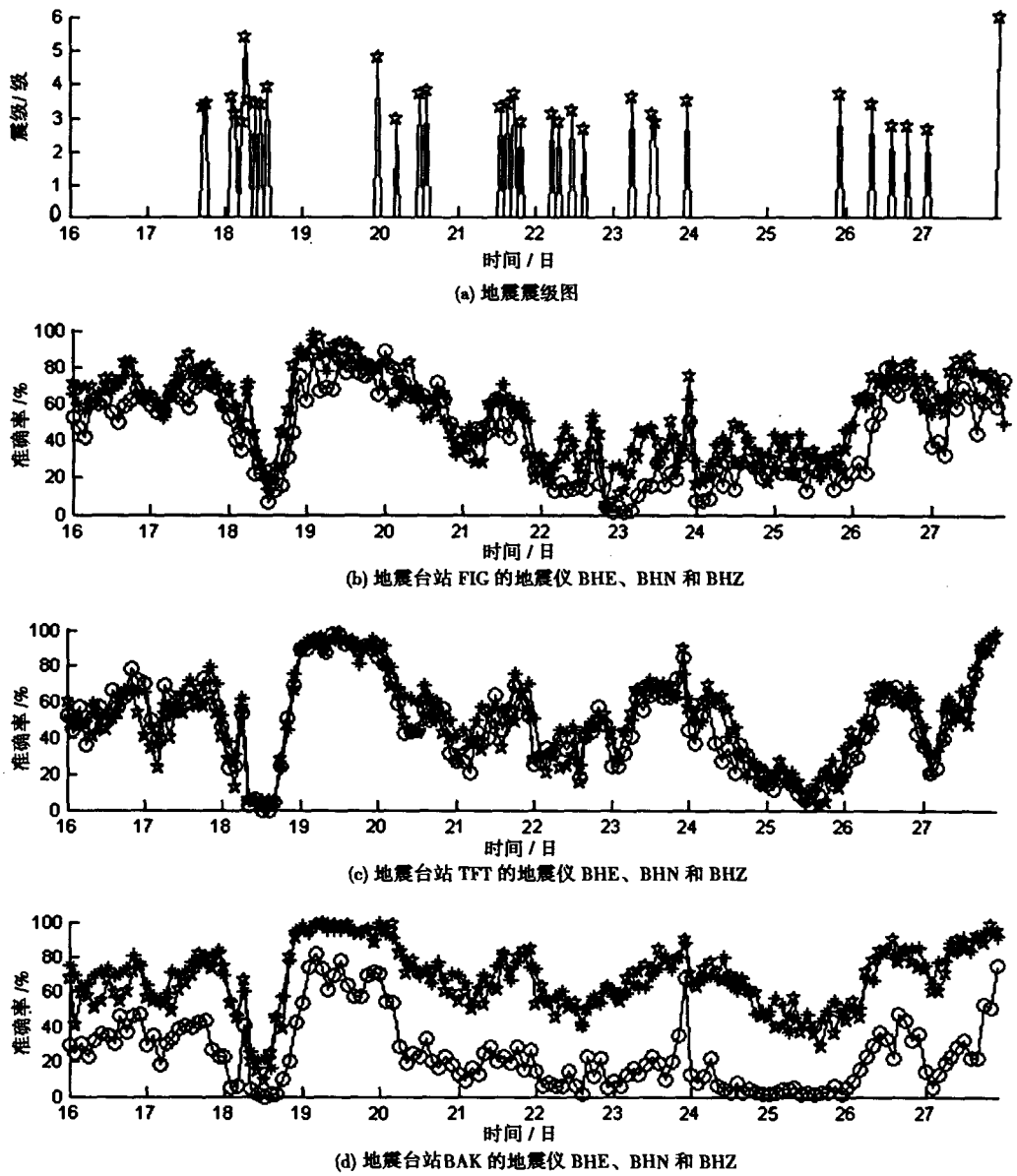


图 4 地震震级及对地震台站 FIG、TFT 和 BAK 的震 2 前数据的预测结果

4 结 论

通过对地震的震前大地异常脉动的提取和分类研究,结果表明 SVM 能有效地将代表地震先兆的异常大地脉动和地震平静时期的大地脉动区分开. 利用震前 SVM 监测模型分析震前大地异常脉动的区域和时间特性,发现在震中断层附近监测到地震震前大地脉动的几率更大. 震前大地异常脉动一般出现在地震前 24 h. 在震前大地异常脉动出现之前,存在一个较短时间的异常平静期. 随着地震的临近,监测到震前大地异常脉动的几率越大. 在离断层较远处,震前监测模型监测到大震的震前大地异常脉动的几率远大于小震. 而震前监测模型在断层附近监测到大震和小震的震前大地异常脉动的几率差别则较小. 这说明小震的震前大地异常脉动同样可以用此模型进行监测.

综上所述,可得:

1)震前 SVM 监测模型能在震前 12 h 内监测到震前大地异常脉动. 随着地震的临近,监测到震前大地异常脉动的几率越高(达 70% ~90%);

2)在震中断层附近监测到震前大地异常脉动的几率高于离震中断层较远的地方;

3)震前 SVM 监测模型监测到大震($M \geq 5$)震前大地异常脉动的几率高于小震($M < 5$).

参考文献：

[1] GELLER R J, JACKSON D D, KAGAN Y Y,et al. Earthquakes cannot be predicted [J]. Science, 1997, 275: 1616.

[2] AKINAGA Y, HAYAKAWA M, LIU J Y, et al. A precursory ULF signature for the Chi-Chi earthquake in Taiwan[J]. Natural Hazards and Earth System Sciences, 2001,1: 33-36.

[3] GOTOH K , HAYAKAWA M , SMIRNOVA N A , et al . Fractal analysis of seismogenic ULF emissions[J]. Physics and Chemistry of the Earth, 2004, 29: 419-424.

[4] IFANTIS A , GIANNAKOPOULOS K . Changes of chaotic behavior of the long-term geoelectric potential difference observed during a five-year investigation and its possible relation to seismic activity in Weatern Greece[J]. Chaos, Solitons and Fractals, 2002, 14: 779-795.

[5] OUILLON G, SORNETTE D. The concept of critical earthquakes applied to mine rockbursts with time-to-failure analysis [J]. Geophysical Journal International, 2000, 143: 454-468.

[6] BIAGI P F, ERMINI A, KINGSLEY S P. Groundwater ion content precursors of strong earthquakes in Kamchatka (Russia) [J]. Pure and Applied Geophysics, 2000, 157: 1359-1377.

[7] MOLCHANOV O , FEDOROV E , SCHEKOTOV A , et al . Lithosphere- atmosphere- ionosphere coupling as governing mechanism for preseismic short- term events in atmosphere and ionosphere[J]. Natural Hazards and Earth System Sciences, 2004, 4: 757-767.

[8] 强祖基，徐秀登，赁常恭. 卫星热红外异常—临震前兆[J]. 科学通报, 1990, 35(17):1324-1327.

[9] TRAMUTOLI V, BELLO G D, PERGOLA N, et al. Robust satellite techniques for remote sensing of seismically active areas[J]. Annali di Geofisica, 2001, 44: 295-312.

[10] TRONIN A A, BIAGI P F, MOLCHANOV O A, et al. Temperature variations related to earthquakes from simultaneous observation at the ground stations and by satellites in kamchatka area[J]. Physics and Chemistry of the Earth, 2004, 29: 501-506.

[11] VAPNIK V. The nature of statistical learning theory[M]. New York: Springer, 1995.

[12] CAI C Z, HAN L Y, JI Z L, et al. SVM-Prot: web-based support vector machine software for functional classification of a protein from its primary sequence[J]. Nucleic Acids Research, 2003, 31(13): 3692-3697.

[13] CAI C Z, HAN L Y, JI Z L, et al. Enzyme family classification by support vector machines[J]. Proteins, 2004, 55(1): 66-76.

[14] CAI C Z, WANG W L, SUN L Z, et al. Protein function prediction via support vector machine approach[J]. Mathematical Biosciences, 2003, 185: 111-122.

[15] CAI C Z, WANG W L, CHEN Y Z. Support vector machine classification of physical and biological datasets[J]. Inter J Mod Phys C, 2003, 14(5): 575-585.

Earthquake Prediction by Using Support Vector Machines

XIAO Han-guang[1,2], *CAI Cong-zhong*[1,2], *YUAN Qian-fei*[1,2], *LIU Xing-hua*[1]

(1. College of Mathematics and Physics, Chongqing University, Chongqing 400030, China;
2. Department of Computational Science, National University of Singapore, Singapore 117543, Singapore)

Abstract: A model of detecting the abnormal earth pulsations was built by the analysis of the earth pulsations before two earthquakes ($M_1 = 6.4$ and $M_2 = 6.0$) took place in the central California of USA via support vector machines. After the analysis and classification of the pre-earthquake earth pulsations recorded by the three nearest earthquake observation stations, it is concluded SVM could differentiate the abnormal earth pulsations from the normal earth pulsations recorded in the quiet phases of earthquake, and the classification accuracy increased with the approach of the two earthquakes. The abnormal earth pulsations appeared 48 and 12 hours before the two earthquakes, respectively. The established model was applied to analysis of the pre-earthquake earth pulsations of the M_1 earthquake (broken out on 22th Dec. 2003) recorded by 14 observation stations in CI earthquake nets. The results showed the model detected the abnormal earth pulsations in the 12 observation stations, and the shorter the distance between observation station and the fault near the epicenter, the higher the probability of detecting the abnormal earth pulsations. This model was also employed to detect the abnormal earth pulsations recorded by three observation stations before the M_2 earthquake. The results revealed the probability of detecting the abnormal earth pulsations ($M \geq 5$) was higher than that of the earthquakes ($M < 5$). This method can be developed to be an effective approach for earthquake prediction.

Key words: support vector machines; earthquake prediction; pre-earthquake seismic-waves; feature extraction

(编辑　李胜春)