

基于粗糙集的 BP 神经网络在震例中的应用研究^{*}

董晓娜¹, 苏道磊², 李希亮¹, 曲利¹, 张慧峰¹, 吴晨¹

(1. 山东省地震局, 山东 济南 250014; 2. 济南市地震局, 山东 济南 250001)

摘要: 采用《中国震例》作为数据源, 通过初步整理分析和预处理, 构建了较完备的震例研究样本集。尝试将粗糙集与 BP 神经网络相结合的方法引入到震例研究中, 用基于粗糙集的属性约简算法从众多复杂的地震异常指标中筛选出对最终分类起决定作用的核心异常作为输入, 震级作为输出, 构建了泛化能力强的 BP 神经网络模型来模拟异常与地震之间的不确定关系。仿真测试结果表明: 地震震级预测精度误差基本控制在 $-0.5 \sim 0.5$ 级之间。

关键词: 粗糙集; 神经网络; 震例研究; 地震异常指标

中图分类号: P315

文献标识码: A

文章编号: 1000-0666(2012)02-0251-09

0 引言

地震预测是世界性科学难题, 因为地震的孕育和发生是很复杂的自然现象。在研究探索中, 人们发现地震发生前会出现大量异常现象, 且异常现象出现的种类多少、持续时间与地震之间有一定的关系, 但这种关系具有很强的不确定性, 是一种非线性映射关系, 很难通过简单的解析表达式来描述, 这使得地震预测具有较高的难度。

神经网络可以通过学习大量样本得到输入与输出之间高度非线性映射关系, 这与一些学者从历史震例中总结出某些规律的地震预测思路相一致。但在实际信息处理时, 一旦输入信息量过大, 神经网络结构就会变复杂, 使得训练时间大大延长, 实效性变差。粗糙集可以解决这个问题, 它可以通过发现数据之间的内在关系、去掉冗余、抽取核心属性从而简化输入。因此, 本文尝试将粗糙集与 BP 神经网络相结合应用于震例研究中, 通过粗糙集属性约简算法对震例数据进行处理, 将众多地震异常中筛选出的核心异常作为输入, 将震级作为输出, 构建泛化能力强的神经网络模型来模拟异常与地震之间的不确定关系, 为地震预测研究提供更客观的指导。

1 理论基础

1.1 粗糙集

1.1.1 概述

粗糙集理论 (Rough Set, 简称 RS), 是波兰学者 Pawlak (1982) 提出的一种能够定量分析处理不精确、不一致、不完整信息与知识的理论方法。从本质上讲, 粗糙集反映了认知过程在非确定性、非模型化信息处理方面的机制和特点, 是一种有效的非单调推理工具。粗糙集和其它处理不确定问题的理论最显著的区别是: 它无需提供任何先验信息, 只需从给定问题的描述集合出发, 找出该问题的内在规律, 所以对问题的描述和处理是比较客观的。

1.1.2 可辨识矩阵

可辨识矩阵 (Discernibility Matrix) 是由波兰华沙大学的著名数学家 Skowron 和 Stepaiuk (1994) 提出来的, 是近年来在粗糙集约简上出现的一个有力工具, 在粗糙集理论中, 它的地位非常重要。其核心思想是: 将数据表中所有与属性区分有关的信息都浓缩在一个矩阵中, 把存在于复杂的信息系统中的全部不可区分关系表达出来, 且不改变原来系统中的潜在知识, 从而大大提高系统获得知识的效率。通过该方法可以很容易求得约简和核, 所以基于可辨识矩阵的属性约简算法简单易

^{*} 收稿日期: 2011-10-09.

基金项目: 中国地震局地震科技星火计划项目 (XH1017) 与山东省地震局合同制项目 (08Q08、09Q14、10Q10) 联合资助。

行,许多属性约简算法都将它作为基础。

设 $S = (U, C \cup D, V, f)$ 为一个决策表, 其中, $U = (x_1, x_2, \cdots, x_n)$ 为论域, $C = \{c_1, c_2, \cdots, c_m\}$ 为条件属性集, $D = \{d\}$ 为决策属性集, 且 $A = C \cup D$,

$$m_{ij} = \begin{cases} 0 & f(x_i, d) = f(x_j, d), \\ \{c_k \mid c_k \in C\} & f(x_i, d) \neq f(x_j, d) \text{ 且 } f(x_i, c_k) \neq f(x_j, c_k), \\ \varnothing & f(x_i, d) \neq f(x_j, d) \text{ 且 } f(x_i, c_k) = f(x_j, c_k). \end{cases} \tag{1}$$

由式 (1) 可知, 当两个样本的决策属性值相同时, 矩阵元素 m_{ij} 定义为 0, 表明这种情况不能区分两个样本; 当两个样本的决策属性值不同时, 分为两种情况: (1) 条件属性值不完全相同时, 矩阵元素 m_{ij} 定义为条件属性值不相同的所有属性集合; (2) 条件属性值完全相同, 说明这两个样本发生冲突, 互不相容, 矩阵元素 m_{ij} 定义为空集。

可辨识矩阵的定义表明: 矩阵中的属性组合数为 1 的元素项是对最终决策起决定作用的, 是决策表的核属性 (核属性可能为空), 是必须保留的。显然, 凡是条件属性组合中包含有核属性的矩阵元素, 都可以仅用核属性就把决策表中决策值不同的记录区分开来, 也就是说对于包含有核属性的矩阵元素而言, 其属性组合中除核属性外的其他条件属性都是多余的, 则这些元素项可被约简。因此, 通过可辨识矩阵可以很容易地得到核属性。

1.1.3 属性约简算法

属性约简是基于粗糙集理论的重要研究内容之一, 范君晖等 (2007)、张运陶和丁保森 (2007)、周必水和黄小克等 (2007)、孙林嘉等 (2009)、徐妙君和吴远红 (2009)、董晓娜等 (2010) 已陆续将基于粗糙集的属性约简算法应用到了工业、商业、社会科学、医学、生物学、网络安全、地震等研究领域。

基于可辨识矩阵的属性约简算法如下:

- 算法 1
- 输入: 决策表 T ; 输出: 约简 RED (初始时, 设置 $RED = \varnothing$)。
- 步骤 1: 计算决策表 T 的可辨识矩阵 M , 并删除可辨识矩阵 M 中的重复项;
- 步骤 2: 依次扫描可辨识矩阵 M 的各元素, 如果存在只含有一个属性的元素, 则记为核属性 CORE ($CORE \in C$), 令 $RED = RED \cup \{CORE\}$, 转至步骤 3。如果不存在只含有一个属性的元素, 转至步骤 5;

V 是属性的值域集, f 是属性空间向值域空间的映射函数 $f: U \times A \rightarrow V$, 则可辨识矩阵可定义为 $M(S) = [m_{ij}]_{n \times n}$ 。其中矩阵项定义为 $m_{ij} (i, j = 1, 2, \cdots, n; k = 1, 2, \cdots, m)$, 公式如下:

步骤 3: 删除可辨识矩阵 M 中包含 CORE 的其他所有元素项;

步骤 4: 如果 $M = \varnothing$, 则输出 RED, 算法结束。否则, 转至步骤 2 继续循环;

步骤 5: 计算剩余属性的重要性 (杨帆, 2005), 并按照其重要性大小进行排序, 选出其中最重要的属性 Csig ($Csig \in C$), 令 $RED = RED \cup \{Csig\}$, 并删除可辨识矩阵 M 中含 Csig 的其他所有元素项;

步骤 6: 如果 $M = \varnothing$, 则输出 RED, 计算结束。否则, 转至步骤 5 继续循环。

1.2 BP 神经网络

1.2.1 人工神经网络

人工神经网络 (Artificial Neural Network, 简称 ANN) 是 20 世纪 80 年代中期兴起的一门科学, 它模拟人脑的一些基本特征, 利用非线性映射和并行处理方法, 通过不断学习和训练, 完成输入与输出之间的映射关系。ANN 在模式识别、机器学习、决策支持、知识发现和数据挖掘等领域都得到广泛的应用。

1.2.2 BP 神经网络

1986 年, 美国加州大学的 Rumelhart 和 McClelland (1986) 在研究并行分布式信息处理方法时, 提出了利用误差反向传播训练算法构造的一种神经网络 Back - Propagation Network, 简称 BP 神经网络。

据统计, 在人工神经网络的应用中, 80% 以上所采用的是 BP 神经网络, 因为其很好地体现了人工神经网络最完美、最精华的部分。它在结构上类似于多层感知器, 是一种有隐含层的多层前馈型网络, 可对非线性可微分函数进行权值训练, 很好地解决了多层网络中隐含单元连接权的学习问题 (Robert, 1987)。由于其具备了结构简单、可调用参数多、训练算法全、可操控性强等优势, 因此得到广泛应用。

具有隐含层 BP 网络的结构如图 1 所示，网络包括输入层、隐含层、输出层。其中隐含层可以由多层组成，每一层有若干个节点，前层和后层通过权值连接，同一层上的各节点之间无耦合连接。BP 网络结构的每一层连接权值都可以通过学习来调节，它的基本处理单元（输入层除外）为非线性的输入输出关系，一般选 Sigmoid 型函数作为传递函数。

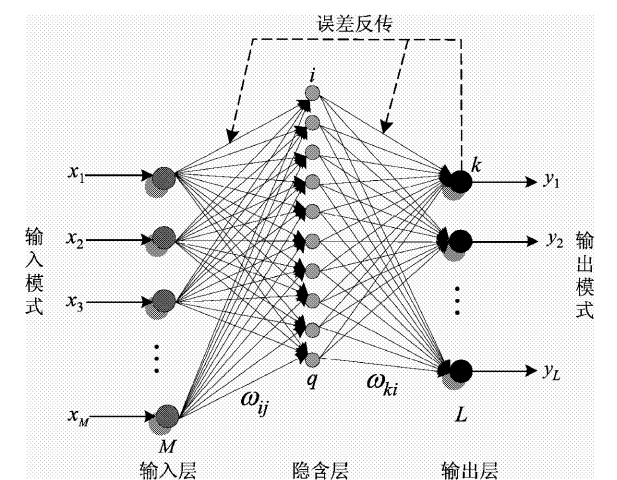


图 1 BP 神经网络结构示意图

Fig. 1 Structure schematic of the BP neural network

2 基于粗糙集的 BP 神经网络在震例中的应用研究

2.1 基于粗糙集的 BP 神经网络应用模型

震前异常的种类多少、持续时间与震级大小

之间的关系是一种非线性映射关系，很难通过简单的解析表达式对其进行描述，如何客观合理地分析这种关系，是地震预测研究领域中的关键问题。神经网络最大的优势是可以通过输入信息充分逼近任意复杂的非线性关系，具有自学习、自适应能力，并有很强的容错性，对输入数据中存在的噪声也不敏感。因此，选用神经网络模型可以很好地模拟地震异常指标与震级之间的关系。但震前出现的异常数量繁多，如果不作处理，将全部指标作为输入来构造神经网络，会使得网络结构复杂，训练时间过长，甚至得不出想要的结果。一些学者凭经验或数理统计挑选一些异常指标构造神经网络，这样有些主观和盲目，对实际的预测结果也可能有一定影响。通过粗糙集属性约简算法提取核心异常指标作为输入，可以降低网络结构的复杂程度，缩短训练时间。因此，本文提出了一种基于粗糙集的 BP 神经网络在震例中的应用模型（图 2）。

基于粗糙集的 BP 神经网络在震例中的应用模型包括粗糙集属性约简模块、神经网络构造及训练模块、神经网络仿真测试模块。各模块的主要功能如下：（1）粗糙集属性约简模块：从震例数据库中提取数据，对其进行异常赋值、震级离散化等相关预处理。将处理后的异常指标构造可辨识矩阵，利用基于可辨识矩阵的属性约简算法进行计算，提取核心异常指标；（2）神经网络构造及训练模块：在 Matlab 环境下，将甄选出的核心异常指标作为输入神经元，离散化震级作为输出神经元，构造 BP 神经网络。调用震例数据库中

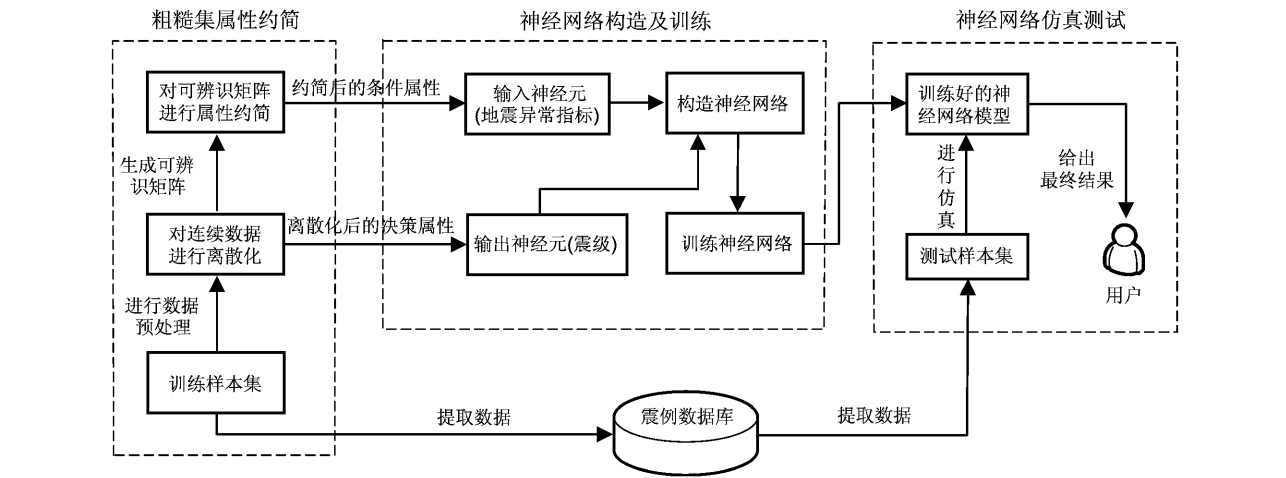


图 2 基于粗糙集的 BP 神经网络在震例中的应用模型

Fig. 2 Application model of the BP neural network based on Rough Set in the earthquake case

75% 的数据作为训练样本集进行训练, 从而建立核心异常指标与震级之间的一种非线性映射关系; (3) 神经网络仿真测试模块: 调用震例数据库中剩余 25% 的数据作为测试样本集, 使用训练好的 BP 神经网络进行仿真测试, 尝试对震级进行预测。

2.2 数据的选取

2.2.1 数据概况

《中国震例》(张肇诚等, 1988, 1990a, b, 1999, 2000; 陈棋福等, 2002a, b, c, 2008) 收录了 1966 ~ 2002 年中国大陆发生的所有 5.0 级以上地震的震例研究报告, 后又增补了 4 条 4.5 ~ 4.9 级的地震震例, 书中对每条震例的地震基本参数及震前异常等都做了详细分析和总结, 是迄今为止国内最完整、最丰富的震例资料, 是地震预测研究和探索地震规律的重要科学资料。在一些学者关于震例的研究中, 已有很多重要的理论成果(付虹等, 2008; 蒋海昆等, 2009)

异常和地震的关系可归纳为 4 种: 1. 异常出现, 有地震发生; 2. 异常未出现, 有地震发生; 3. 异常出现, 无地震发生; 4. 异常未出现, 也无

地震发生。这 4 种关系可以描述为以下 4 种情况: A. 从异常出现层面来说, 1 和 3 构成理论上的完备集; B. 从地震发生层面来说, 1 和 2 构成理论上的完备集; C. 从异常未出现层面来说, 2 和 4 构成理论上的完备集; D. 从无地震发生层面来说, 3 和 4 构成理论上的完备集。其中, 4 无有用信息, 包含 4 的 C 和 D 在实际研究中价值不大。2 和 3 的情况很常见, 也正是造成大量“漏报”(报无实有)和“虚报”(报有实无)的原因(吴绍春, 2005)。完整意义下的研究应该是 A 和 B 的总和, 但是由于 3 的缺失或不完备使得 A 成了实际上的不完备集, 所以无法开展研究。因此本文的研究只针对 B。

2.2.2 数据选取

本文对《中国震例》中记载的 209 条震例的异常信息进行详细整理, 按观测手段统计: 测震学异常指标共计 74 项, 前兆异常指标共计 84 项。经整理分析和统计归纳, 最终得到异常指标共 106 项, 其中测震学指标 41 项, 前兆指标 65 项, 详细信息如表 1 所示。

表 1 地震异常指标详表
Tab. 1 Detailed table of seismic anomaly indicators

异常分类	异常指标序号及名称
测震学指标	1. 地(孕)震空区; 2. 背景空区; 3. 诱发地震(响应地震、地震丛集); 4. 前兆震(群); 5. 相关地震; 6. 前震活动; 7. 地震节律; 8. 应变释放; 9. 能量释放; 10. 地震窗口; 11. 地震频度; 12. 地震条带; 13. 震群活动; 14. 缺震; 15. 地震时空分布; 16. 地震迁移; 17. 地震平静; 18. 地震集中; 19. 地震活动增强; 20. 小震调制比 R_m ; 21. 断层面总面积 $\sum(t)$ 值; 22. 地震活动性指标(指数 A 值、标度 ΔF 、活动熵 Q 值、活动度 S 值、模糊地震活动度 S_y 值); 23. 三项指标: 频度、面积、震级 ($N \setminus S \setminus E$); 24. 震级容量维 D_0 (分形); 25. b 值; 26. 地震空间集中度 C 值; 27. 地震危险度 D 值; 28. 算法复杂性 $A_c(C_n)$ 值; 29. 震情指数 $A(b)$ 值; 30. 地震时间间隔 dt (复发周期); 31. η 值 ($G-R$ 关系偏离度); 32. h 值; 33. 地震强度因子 M_f 值; 34. GL 值 (地震非均匀度); 35. 尾波 (尾波 q 值、尾波持续时间比 τ 、衰减系数 α); 36. S 波偏震; 37. 波速比; 38. 地脉动; 39. 振幅比 A_s/A_p ; 40. 其他震源及介质参数 (应力降、介质因子 Q 值、剪切应力值); 41. P 波初动解 (P 波初动符号矛盾比、 P 轴偏转、小震综合节面解)
	1. (定点) 重力; 2. 流动重力; 3. 短水准; 4. 长水准; 5. 流动水准; 6. 基线; 7. 短基线; 8. 流动基线; 9. 地倾斜; 10. 水管倾斜; 11. 水平摆倾斜; 12. 竖直摆倾斜; 13. 伸缩仪; 14. 跨断层蠕变; 15. 振弦应变; 16. 钻孔应变; 17. 压容应变; 18. 电感地应力; 19. (视) 电阻率; 20. 地电; 21. 自然电位; 22. 电磁波; 23. 流动地磁; 24. 地磁 (地磁转换函数); 25. 地磁低点位移; 26. 地磁垂直分量 (Z); 27. 地磁偏角 (D); 28. 地磁总场 (F); 29. 水位; 30. 气压; 31. 气温; 32. 地温; 33. 水温 (温泉水温); 34. 温泉 (泉水) 流量; 35. 地下水流量; 36. pH 值; 37. 电导率; 38. 水导电; 39. 气体总量; 40. 二氧化碳 (CO_2); 41. 氢气 (H_2); 42. 氮气 (N_2); 43. 气氦 (Rn); 44. 水氦; 45. 气汞 (Hg); 46. 水汞 (Hg); 47. 甲烷 (CH_4); 48. 氧+氩 ($O_2 + Ar$); 49. 水化 (硫化氢 H_2S); 50. 氦 (He); 51. 氟离子 (F^-); 52. 氯离子 (Cl^-); 53. 碳酸氢根离子 (HCO_3^-); 54. 重碳酸离子 (CO_3^{2-}); 55. 硫酸根离子 (SO_4^{2-}); 56. 钙离子 (Ca^{2+}); 57. 镁离子 (Mg^{2+}); 58. 锂离子 (Li^+); 59. 钾离子 (K^+); 60. 二氧化硅 (SiO_2); 61. 前兆信息熵; 62. 旱涝; 63. 海平面; 64. 宏观; 65. 石油井动态

2.3 数据预处理

2.3.1 异常赋值

异常种类多少及其持续时间是研究问题的两个关键因素。异常的数量，可直接相加求得，但异常持续时间则无法直接获取。本文根据每条震例中异常的出现情况，通过统计其起止时间进行赋值。以 SAI_i (Seismic Anomaly Indicator) 来表示第 i 项异常指标，用 T_i 来表示第 i ($i=1,2,\cdots,106$) 项异常指标的持续时间，单位为月。分以下3种情况来讨论：

(1) 异常未出现：如果某个地震之前，某项异常未出现，也就是书中未记载该项异常信息，那么其对应的 T 赋值为0。例如：“1975年2月4日辽宁海城7.3级地震”，未出现指标“应变释放” SAI_8 异常，因此，笔者给 SAI_8 赋值为 $T_8=0$ 。

(2) 某项异常只在单测点出现：某次地震之前，只有一个测点出现异常， T 取该异常的持续时间。例如：“1976年4月6日内蒙古和林格尔6.3级地震”，1962年5月至1973年4月在和林格尔地区 (300×400) km^2 范围内应变释放出现加速密集异常现象，起止时间共计132个月。因此，笔者给第15号震例的属性 SAI_8 “应变释放” 赋值为 $T_8=132$ 。

(3) 某项异常同时出现在多个测点：某次地震之前，也可能存在某项异常同时在多个测点出现的情况。对于这种情况，异常持续时间如何赋值是个关键问题，因为在构造BP网络时，要把异常持续时间作为输入，如果处理不得当，会直接影响BP神经网络的结构，从而得不到理想的结果(王炜等，2000；刘悦，2008)。本文采用加权求和法来解决以上问题，即把该异常出现在各测点的持续时间进行加权求和。假设某次地震前，某项异常指标 SAI_i ($i=1,2,\cdots,106$) 在多个测点 j ($j=1,2,\cdots,n$) 均出现，则 SAI_i 的持续时间 T_i 可定义为

$$T_i = \sum_{j=1}^n \omega_j t_{ij} \tag{2}$$

其中， t_{ij} 表示第 i 项异常指标在第 j 个测点的持续时间， ω_j 为权值，定义如下：

$$\omega_j = t_{ij} / \sum_{j=1}^n t_{ij} \tag{3}$$

将式(3)带入式(2)，可得

$$T_i = \sum_{j=1}^n (t_{ij} / \sum_{j=1}^n t_{ij}) t_{ij} = \frac{\sum_{j=1}^n (t_{ij})^2}{\sum_{j=1}^n t_{ij}} \tag{4}$$

例如：“1973年2月6日四川炉霍7.6级地震”的“地震频度”指标 (SAI_{11}) 在两个测点均出现了异常，分别持续了2个月和20个月。因此，根据式(4)，可计算该震例的异常 SAI_{11} 持续时间为 $T_{11}=18.36$ 。

2.3.2 震级离散化

由于粗糙集方法不能直接处理连续属性，因此需要先把震级进行离散化处理，本文选择等距离离散法将震级划分，如表2所示。

表2 震级离散化
Tab.2 Magnitude discretization

原始震级 M_L	离散化后的震级
4.5 ~ 5.4	5
5.5 ~ 6.4	6
6.5 ~ 7.4	7
7.5 ~ 8.4	8
8.5 ~ 9.4	9

2.4 基于粗糙集的属性约简

2.4.1 构造可辨识矩阵

(1) 论域

由于异常指标共106项，数据量繁多，如果直接进行处理，可能会降低约简算法的执行效率。鉴于测震和前兆是相对独立的研究体系，本文将震例数据分为测震学指标和前兆指标，并分别构造可辨识矩阵来进行约简。去除测震学指标中震前无任何测震异常的7条震例，剩余的202条震例信息构成研究测震学指标的论域。同理，对于前兆指标，去除震前无任何前兆异常的17条震例，剩余的192条震例信息构成研究前兆指标的论域。

(2) 条件属性

《中国震例》记载的每一条震例中，很多异常指标可能是震后通过研究认识到它与地震的关系后加上去的。根据目前的资料已无法分辨哪些异常是震前确实存在，并且被地震工作者识别到的；哪些异常是震前确实出现了，但无法断定是否是异常；哪些异常是震后通过总结追加上去的。因此本文约定：无论哪种情况，凡是在某条震例记

载中出现的异常指标,在决策表的条件属性中都表示为 1,未在震例记载中出现的则为 0。

(3) 决策属性

如果决策属性以“地震是否发生”来赋值,每一条震例的决策属性都是 1,就起不到决策作用。考虑到本文研究的是震前异常出现的种类多少、持续时间与震级的关系,所以选择震级的大小作为决策属性。

综上所述,就测震学指标而言,论域为 202 个

震例,条件属性为 41 个测震异常,决策属性为震级。根据条件属性的取值,每条震例可描述为一个由 41 个“0 或 1”组成的序列。那么,202 个震例就对应 202 个序列,可根据式(1)生成测震学指标的可辨识矩阵,即 (202×202) 阶的矩阵,如表 3 所示。同理,对于前兆指标,论域为 192 个震例,条件属性为 65 个前兆异常,决策属性为震级,也可按上述原理生成前兆指标的可辨识矩阵,即 (192×192) 阶的矩阵。

表 3 测震学异常指标的可辨识矩阵
Tab. 3 Discernibility matrix of seismometry anomaly indicators

	$i=1$	$i=2$	$i=3$	$i=4$	$i=5$	$i=202$
$j=1$	0	0	06	01 06 12	01 06 11 12 16 37	01 06 11 14 22 25 29 32 39 40 41
.....
$j=202$	01 06 11 14 22 25 29 32 39 40 41	01 02 10 12 14 22 29 32 39 40 41	0	11 12 14 22 25 29 32 39 40 41	0	0

2.4.2 基于可辨识矩阵的属性约简

根据算法 1,逐项扫描可辨识矩阵,对于重复项只保留其中一项。检索可辨识矩阵中只含有一个属性的项即核属性 Core,将其加入约简 RED,并删除可辨识矩阵中含 Core 的所有项。若可辨识矩阵已为空,则 RED 就为最终所求的约简,若不为空,则计算剩余属性的重要性,依重要性排序,依次将目前最重要的属性 Csig 加入到 RED,并删减掉包含 Csig 的所有项,直到矩阵为空为止。

对测震学指标的可辨识矩阵进行约简,最终得出测震学指标的核属性依次为:06、11、10、02、08、25、16、37、12、04、01、15、14、20、39、09、21、17、34、19、31、22。然后将包含这些核属性的矩阵项约简后,矩阵为空,算法结束,以上的核属性即为所求的约简 RED。整理可得经过粗糙集属性约简后的核心测震学异常指标如表 4

所示。

同理可对前兆指标的可辨识矩阵进行分析,得到的核属性依次为:33、03、64、09、44、18、19、29、16、10、02、11、01、25、27、17、28,将它们存入 RED,并将包含这些核属性的矩阵项约简后,矩阵仍不为空。然后,计算剩余属性的重要度 Csig,依重要度递减排序,按照排序继续进行约简,直到矩阵为空,该算法结束。又加入到 RED 的属性有:22、26、43、38、21、49、31。整理可得经过粗糙集属性约简后的核心前兆异常指标如表 5 所示。由表 4、表 5 可知,使用基于粗糙集属性约简的方法后,最终得出 46 项核心指标,其中测震学分类 22 项,前兆分类 24 项。整理所有经过粗糙集属性约简后的震例信息,即对每条震例记录中冗余信息进行删减,只保留 46 项核心异常指标。

表 4 约简后的测震学异常核心指标
Tab. 4 Core indicators of seismometry anomaly after reduction

异常分类	异常指标序号及名称
测震学指标	1. 地 (孕) 震空区; 2. 背景空区; 3. 前兆震 (群); 4. 前震活动; 5. 应变释放; 6. 能量释放; 7. 地震窗口; 8. 地震频度; 9. 地震条带; 10. 缺震; 11. 地震时空分布; 12. 地震迁移; 13. 地震平静 14. 地震活动增强; 15. 小震调制比 R_m ; 16. 断层面总面积 $\sum(t)$ 值; 17. 地震活动性指标 (指数 A 值; 标度 ΔF ; 活动熵 Q 值; 活动度 S 值; 模糊地震活动度 S_y 值); 18. b 值; 19. η 值 ($G-R$ 关系偏离度); 20. GL 值 (地震非均匀度); 21. 波速比; 22. 振幅比 A_S/A_P

表 5 约简后的前兆异常核心指标

Tab. 5 Core indicators of the precursory anomaly after reduction

异常分类	异常指标序号及名称
前兆指标	1. (定点) 重力; 2. 流动重力; 3. 短水准; 4. 地倾斜; 5. 水管倾斜; 6. 水平摆倾斜; 7. 钻孔应变; 8. 压容应变; 9. 电感地应力; 10. 自然电位; 11. (视) 电阻率; 12. 电磁波; 13. 地磁低点位移; 14. 地磁垂直分量 (Z); 15. 地磁偏角 (D); 16. 地磁总场 (F); 17. 水位; 18. 气温; 19. 水温 (温泉水温); 20. 水导电; 21. 气氦; 22. 水氦; 23. 水化 (硫化氢 H ₂ S); 24. 宏观

2.5 BP 神经网络构造训练及仿真测试

(1) 样本选择

调用经过粗糙集属性约简得到的核心属性的震例数据，将其中的 162 条震例（约占总样本数的 75%）作为训练样本，其余的 40 条震例（约占总样本数的 25%）作为测试样本，在 MATLAB R2007a 环境下构造 BP 网络模型。

(2) 归一化与反归一化

由于隐含层中的神经元选用 Sigmoid 型传递函数，所以在构造神经网络之前，先将训练样本归一化，使各样本元素的值为 [0, 1]，这样才能确保网络对样本具有足够的输入敏感性和良好的拟合性。训练好的 BP 网络模型对测试样本进行仿真之后，需对仿真结果进行反归一化得到实际的震级。本文选用的归一化与反归一化方法如式 (5) 和式 (6) 所示。

$$\hat{X} = \frac{X - X_{\min}}{X_{\max} - X_{\min}}, \tag{5}$$

$$X = X_{\min} + \hat{X}(X_{\max} - X_{\min}). \tag{6}$$

(3) BP 神经网络的构造、训练及仿真测试

神经网络构造的主要工作是确定神经网络的层数和各层的神经元数目。输入神经元使用粗糙集属性约简后得出的 46 项核心指标，包括测震指标和前兆指标；输出神经元即震级大小。

将训练样本以矩阵的形式导入神经网络模型，本文训练样本集共有 162 个样本，因此，该神经网络的输入为一个 (46 × 162) 阶矩阵，输出为一个 (1 × 162) 阶矩阵。在 MATLAB R2007a 环境下，建立了 BP 神经网络模型进行训练，如图 3 所示。

本文中测试样本集共有 40 个样本，因此，输入为一个 (46 × 40) 阶矩阵，输出为一个 (1 × 40) 阶矩阵。在 Matlab 环境下，将测试样本以矩阵的形式导入训练好的 BP 神经网络模型进行仿真测试。

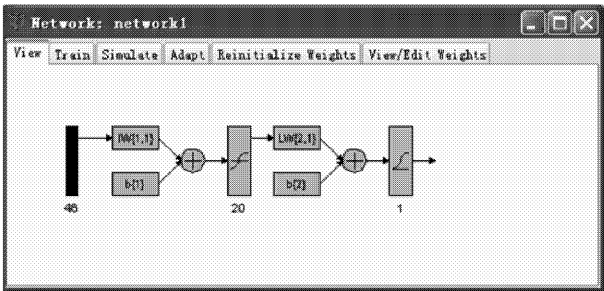


图 3 MATLAB 下建立 BP 网络模型

Fig. 3 BP neural network model based on MATLAB software

(4) 实验结果及分析

经过反复实验并对比多次实验的结果，得到以下结论：当隐含层神经元数设为 20，训练误差设为 0.01 时，训练得到的 BP 神经网络结构最佳。用测试样本进行仿真实验，误差结果如图 4 所示。用来仿真实验的 40 个测试样本中，除了 2 个样本的震级差略大（图中以圆圈表示），其余样本的震级差都小于 0.5（图中虚线之间部分）。

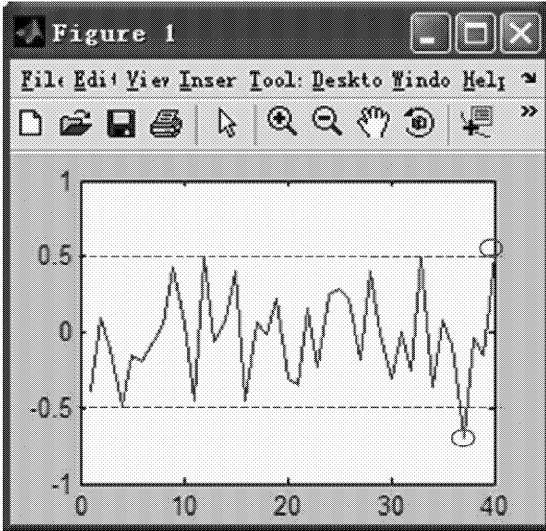


图 4 BP 神经网络训练及仿真结果图

Fig. 4 Training and simulation results of BP neural network

本文将基于粗糙集的 BP 神经网络的仿真实验误差与陈以等 (2011) 提出的组合人工神经网络的研究结果进行对比, 结果如表 6 所示。可见, 无论是最大误差还是预测误差, 基于粗糙集的 BP 神经网络预测效果都明显优于传统的 BP 神经网络。其中, 最大误差减少了 36.92%, 平均误差减少了 17.02%, 在最大误差上, 基于粗糙集的 BP 神经网络的预测效果更占优势, 最大误差减少了 0.080 7, 但平均误差却增大 0.047 4。因此, 以上两种研究方法各有优劣, 可在实际的应用中根据需求来选择。

表 6 误差结果对比
Tab. 6 Contrast of error results

	传统 BP 神经网络	组合人工 神经网络	基于 RS 的 BP 神经网络
最大误差	0.993 3	0.707 3	0.626 6
平均误差	0.373 5	0.242 5	0.289 9

3 结论与探讨

前人使用神经网络进行地震预测的研究中存在着以下两个问题:

(1) 样本量小。蔡煜东等 (1993)、杨居义和易永宏 (2008)、陈一超等 (2008)、陈以等 (2011) 仅选用部分区域的部分震例进行研究。如: 选用云南滇西南地震区的 17 条震例资料作为研究样本。再如: 王伟等 (2000) 选用大华北地区 1975 年至 1998 年 7 月发生 5 级以上的 32 条震例资料作为研究样本。选择小量样本进行研究, 可能会导致研究结果以面概全, 不够客观和全面。

(2) 人为选择异常指标作为输入。蔡煜东等 (1993)、王伟等 (2000)、杨居义和易永宏 (2008)、陈一超等 (2008)、陈以等 (2011) 通常根据经验或简单的数理统计, 对神经网络模型的输入人为进行选择。这种做法可能会使得整个研究过程不够客观和全面。

本文选用目前能够收集到的最完整的震例数据作为研究样本, 以便在大量震例研究的基础上提取共性特征, 使得研究结果更具普适性。然后提出了一种基于粗糙集的神经网络在震例中的应用模型, 实现了用粗糙集属性约简算法来自动筛选核心地震异常指标, 最后通过 BP 神经网络建立

核心异常指标与地震震级之间的一种非线性映射关系, 尝试对地震震级进行预测。实验结果如下:

(1) 对《中国震例》中记载的 209 条震例的异常信息做了详细整理, 最终得到异常指标共 106 项, 其中测震学指标 41 项, 前兆指标 65 项, 构成了进行震例研究的较完备样本集。

(2) 使用基于粗糙集属性约简的方法, 从 106 项异常指标筛选出了 46 项核心指标, 其中测震异常 22 项, 前兆异常 24 项。这 46 项中囊括了前人经验研究中常使用的异常指标, 如测震学异常指标中的地震条带、应变释放、 b 值和波速比等, 前兆异常指标中的短水准、地倾斜、电位、水氡、水位等。这些指标可以从众多指标中通过粗糙集属性约简算法被自动筛选出来, 足以证明该算法在解决此问题上的科学性和正确性。另外, 该方法还筛选出了: 前震活动、地震迁移、地震平静、地震活动增强等有关地震时空分布方面的属性; 以及小震调制比 R_m 、断层面总面积 $\sum(t)$ 值、地震活动性指标、 η 值、 GL 值、振幅比等异常指标; 形变学科中的重力、钻孔应变, 电磁学科的电阻率、地磁低点位移、地磁 Z 、 D 、 F , 以及流体学科的气温、水温、硫化氢等异常指标。根据粗糙集属性约简算法的特性, 得出的核心属性都是对最终分类起决定作用的。因此在今后的研究中, 这部分异常指标的使用值得引起足够重视, 将它们合理地应用到实际的地震预测研究中, 可能会得到更客观、更全面的结果。

(3) 最终预测的震级与实际的震级误差基本控制在 $-0.5 \sim 0.5$ (即不超过 0.5 级), 结果基本符合地震预测精度要求。因此, 使用粗糙集与神经网络相结合的方法对震级进行预测研究是行之有效的。今后, 还可尝试使用数据挖掘中的聚类分析、关联规则、支持向量机等算法对核心异常开展进一步的探索研究。

本文的撰写得到山东师范大学信息科学与工程学院段会川教授的悉心指导, 在此表示诚挚感谢!

参考文献:

蔡煜东, 甘俊人, 姚林声. 1993. 人工神经网络在地震综合预报中的应用[J]. 地震学报, 15(2): 257-260.
陈棋福, 郑大林, 车时, 等. 2008. 中国震例(2000-2002)[M]. 北京: 地震出版社.

- 陈棋福,郑大林,车时. 2002a. 中国震例(1992-1994)[M]. 北京:地震出版社.
- 陈棋福,郑大林,高荣胜. 2002b. 中国震例(1997-1999)[M]. 北京:地震出版社.
- 陈棋福,郑大林,刘桂萍,等. 2002c. 中国震例(1995-1996)[M]. 北京:地震出版社.
- 陈一超,曾三友,张好春,等. 2008. 基于遗传神经网络的地震预测研究[J]. 计算机应用与软件,25(4):135-137.
- 陈以,王颖. 张晋魁. 2011. 组合人工神经网络在地震预测中的应用研究[J]. 计算机仿真,28(1):190-193.
- 董晓娜,段会川. 2010. 基于粗糙集的属性约简在地震异常指标识别中的应用研究[J]. 山西地震,(1):21-24.
- 董晓娜. 2010. 粗糙集支持的人工神经网络在地震预测中的应用研究[D]. 济南:山东师范大学.
- 范君晖,吴忠,李旭芳. 2007. 基于粗糙集和神经网络的上海最低工资标准研究[J]. 计算机工程与设计,28(12):2 918-2 921.
- 付虹,李永莉,赵小艳,等. 2008. 云南 $M \geq 5$ 地震震前异常的统计特征[J]. 地震研究,31(4):335-339.
- 蒋海昆,苗青壮,吴琼,等. 2009. 基于震例的前兆统计特征分析[J]. 地震学报,31(3):245-259.
- 刘悦. 2005. 神经网络集成及其在地震预报中的应用研究[D]. 上海:上海大学.
- 孙林嘉,李茹,屈元子. 2009. 煤矿瓦斯预测知识获取模型的应用研究[J]. 计算机工程,35(12):169-171.
- 王炜,吴耿锋,宋先月. 2000. 神经网络在地震学方法综合预报中的应用[J]. 地震学报,22(3):189-193.
- 吴绍春. 2005. 地震预报中的数据挖掘方法研究[D]. 上海:上海大学.
- 徐妙君,吴远红. 2009. 基于粗糙集和神经网络结合的鱼病诊断方法[J]. 计算机工程与设计,30(7):1738-1741.
- 杨帆. 2005. 粗糙集约简算法及其应用的研究[D]. 武汉:武汉科技大学.
- 杨居义,易永宏. 2008. 基于 BP 神经网络的地震预测研究[J]. 微电子学与计算机,25(10):129-132.
- 张运陶,丁保森. 2007. RS-GA-SVM 方法在真核生物 DNA 翻译起始位点识别中的应用[J]. 计算机与应华学,6(24):794-798.
- 张肇诚,罗兰格,李海华,等. 1988. 中国震例(1966-1975)[M]. 北京:地震出版社.
- 张肇诚,罗兰格,李海华,等. 1990a. 中国震例(1976-1980)[M]. 北京:地震出版社.
- 张肇诚,罗兰格,李海华,等. 1990b. 中国震例(1981-1985)[M]. 北京:地震出版社.
- 张肇诚,郑大林,徐京华. 1999. 中国震例(1986-1988)[M]. 北京:地震出版社.
- 张肇诚,郑大林,徐京华. 2000. 中国震例(1989-1991)[M]. 北京:地震出版社.
- 周必水,黄小克. 2008. 基于粗糙集的案例推理在网络行为安全审计监控中的应用[J]. 计算机应用与软件,3(25):249-251.
- Pawlak Z. 1982. Rough sets [J]. International Journal of Information and Computer Science,11(5):341-356.
- Robert H N. 1987. Counter propagation Networks[C]. USA:San Diego.
- Rumelhart D E, McClelland J L. 1986. Parallel Distributed Processing: Explorations in the Microstructure of Cognition [M]. Cambridge, MA:MIT Press.
- Skowron A, Stepiauk J. 1994. Decision Rules Based on Descrinability Matrices and Decision Matrices[C]. California:San Jose,602-609.

Application of BP Neural Network Based on Rough Set in the Earthquake Case

DONG Xiao-na¹, SU Dao-lei², LI Xi-liang¹, QU Li¹, ZHANG Hui-feng¹, WU Chen¹

(1. Earthquake Administration of Shandong Province, Jinan 250014, Shandong, China)

(2. Earthquake Administration of Jinan Municipality, Jinan 250001, Shandong, China)

Abstract

Firstly, using "China earthquake case" as data source, we built a fairly complete sample set for earthquake case study through preliminary analysis and pretreatment and introduced the combination of Rough Set and BP Neural Network to the earthquake case. Secondly, we selected the core abnormalities which plays a decisive role in final classification from a number of complex seismic anomaly indicators as the input by use of attribute reduction algorithm based on Rough Set, and took the discrete magnitude as the output. Furthermore, we built a generalized BP Neural Network model to simulate the uncertain relationship between the seismic anomaly and the earthquake. Finally, the result of simulation tests showed that the precision errors of earthquake magnitude prediction is between -0.5 and 0.5.

Key words: Rough Set; Neural Network; study of earthquake cases; seismic anomaly indicator