

粗糙神经网络在建筑物震害预测中的应用

刘勇健^{1,3}, 刘义建², 张伯友³

(1. 广东工业大学, 广东 广州 510006; 2. 湖南省地质矿产勘查开发局 407 队, 湖南 怀化 418000;
3. 中国科学院广州地球化学所, 广东 广州 510640)

摘要:将粗糙集理论和神经网络原理结合起来,建立了基于粗糙-神经网络的建筑物震害预测模型。首先运用粗糙集理论,根据原始样本建立决策表进行属性离散化、属性重要性排序、属性约简和分类规则的提取;然后将所提取的关键成分作为神经网络的输入训练模型。实例研究表明,基于粗糙-神经网络的多层砖房震害预测结果与实际震害基本吻合。该模型简化了神经网络结构,提高了训练速度和分类精度,还能对各因素对房屋震害的影响度进行分析。

关键词:粗糙集;神经网络;震害预测;属性约简

中图分类号: P315.9

文献标识码: A

文章编号: 1000-0844(2008)02-0097-05

Application of Seismic Damage Prediction to Buildings Based on Rough Set and Neural Network

LIU Yong-jian^{1,3}, LIU Yi-jian², ZHANG Bo-you³

(1. Guangdong University of Technology, Guangzhou 510006, China;
2. 407 Brigade, Geology and Mineral Resources Exploration and Development Bureau of Hunan Province, Hunan Huaihua 41800, China;
3. Guangzhou Institute of Geochemistry Chinese Academy of Sciences, Guangzhou 510640, China)

Abstract: Rough set theory and artificial neural network are integrated into a model of seismic damage prediction for buildings. First the rough set theory is used to acquire the knowledge of classification, which includes the decision table construction, attribute discretization, attribute importance ranking, attribution reduction and rule abstract. Then the key components are extracted as the input of the neural network. The method reduces the structure of neural network model, and raises efficiency of training and accuracy of prediction. The importance ranking of these factors to earthquake-resistance performance can be obtained by this model. The research shows that the prediction results agree with actual seismic damage of multistory masonry building.

Key words: Rough set theory; Neural network; Seismic damage prediction; Attribution reduction

0 引言

我国为地震多发国家,据统计,20世纪100年间6级以上的破坏性地震在我国共发生650次余,死亡约59万人;全球死亡20万人以上的两次地震均发生在我国;我国有67%的大城市位于地震烈度Ⅶ度及以下的地区。因此建筑物震害预测是城市防震减灾工作的重要组成部分。为了避免或减轻地震灾害的损失,国内外许多专家学者在震害预测方面

做了大量研究,提出了多种震害预测方法,如历史地震统计法、半经验半理论法、模糊类比法、结构计算法等^[1-9]。震害研究表明,建筑物震害一方面与地震特性相关,另一方面又与建筑结构特性有关。由于各影响因素的随机性、模糊性及统计资料的不完整性,震害预测系统是一个高度复杂非线性系统。

人工神经网络(artificial neural network, ANN)

收稿日期:2007-12-03

基金项目:广东省自然科学基金项目(6021462)

作者简介:刘勇健(1968-),女(汉族),湖南安化人,博士生,副教授,主要从事岩土工程、地质工程的教学与研究工作。

是20世纪80年代中期兴起的一门非线性科学,它能处理各种模糊的、非线性的、不精确并带有噪声的数据,具有很强的学习功能和容错能力,已在模式识别、数据处理、自动控制等领域得到广泛的应用。粗糙集理论(Rough Set, RS)是波兰数学家 Z. Pawlak 于1982年提出的一种数据分析理论,已成为处理模糊和不确定知识的新型数学工具。本文将粗糙集理论和神经网络原理有机地结合起来,建立了基于粗糙集-神经网络(Rough Neural Networks, RNN)的建筑物震害预测模型,为建筑物震害预测尝试了一种新的研究途径。

1 粗糙集-神经网络的基本理论

1.1 粗糙集理论(RS)

粗糙集方法把研究对象视作一个信息系统,通过不可分辨关系和不可分辨类确定给定问题的近似域,从而找到问题的内在规律。根据知识表达中不同属性的重要性,明确有用信息和冗余信息,简化输入信息的表达空间维数。在保持分类能力不变的前提下,通过知识约简,直接从训练数据中导出分类规则。

定义1(决策系统) 一个信息系统表示为 $S = (U, R, V, f)$ 。其中, U 为对象的非空有限集合,即论域; R 为属性的非空有限集合; $V = \bigcup_{r \in R} V_r$, V_r 为属性 r 的值域; $f: U \times R \rightarrow V$ 是一个信息函数,指定 U 中各对象的属性唯一值。对信息系统 S ,若属性集 $R = C \cup D$,子集 C 和 D 分别称为条件属性集和决策属性集, $C \cap D \neq \emptyset$,则该信息系统称为一个决策系统或决策表。如表1。

表1 决策表的一般表达形式

对象	条件属性 C			决策属性 D
U	C_1	...	C_n	D
U_1	C_{11}	...	C_{1n}	d_1
U_2	C_{21}	...	C_{2n}	d_2
\vdots	\vdots	\vdots	\vdots	\vdots
U_m	C_{m1}	...	C_{mn}	d_m

1.2 人工神经网络

人工神经网络是一种模仿人的大脑结构和功能

的信息处理系统,是由大量神经元广泛连接而成的复杂网络。它通过样本的学习、神经元的模拟、记忆和联想建立输入空间到输出空间的非线性映射关系。近年来应用最广泛的是BP网络(Back propagation Neural Networks, BP),最基本的BP网络由输入层、隐含层和输出层组成。BP学习过程是由两部分组成:即信息的正向传递与误差的反向传播。在正向传播过程中,输入信号从输入层经隐含层向输出层逐层传递,并计算各单元的实际输出值;每一层神经元的输出作用于下一层神经元的输入,若在输出层不能得到期望的输出,则将实际输出与目标输出之差值(即误差)反向传播,通过网络将误差信号沿原来的连接通路反传回来修改各层的权值,反复调整,直至误差信号达到期望目标^[10]。根据Kolmogorov定理,在一定条件下对于任意给定的 $\varepsilon > 0$ 存在一个3层神经网络,能以 ε 均方差的精度逼近任意连续函数,因此可用于建筑物震害预测。

RS对数据的分析能力很强,但对噪声较敏感;ANN具有很强的容错性和并行处理能力,但不能确定哪些知识是冗余的,哪些知识是有用的,并且当输入信息空间的维数较大时,网络结构复杂,训练时间过长,制约其实用化。本文将两者有机结合起来,建立基于粗糙集理论框架下的神经网络预测模型。

2 基于粗糙集-神经网络(RNN)的建筑物震害预测模型

2.1 震害影响因子分析

城市中虽然新建了大量高层建筑,但还存在一定比例老旧房屋和一些抗震性能较差的房屋,其中多层砖房问题尤为突出。本文对多层砖房在确定的地震强度作用下,发生某种破坏程度的概率或可能性进行预测研究。多层砖房在地震作用下破坏的程度可划分为五个等级:基本完好、轻微破坏、中等破坏、严重破坏、毁坏,如表2所示。

大量的宏观震害经验表明,下列因素与多层砖房的震害程度密切相关^[3,4]。

表2 多层砖房破坏等级的划分标准

破坏状态	代号	震害描述
基本完好	I	墙体完好,少数非主体结构有局部轻微破坏,或个别突出部分偶有轻微破坏裂纹
轻微破坏	II	非主体结构局部破坏有明显的破坏,或少数墙板有轻微裂纹,少数墙体偶有明显裂纹
中等破坏	III	主体结构或连接部分多处发生明显的裂缝,或填充墙体、附属建筑等严重破坏,甚至倒塌
严重破坏	IV	主体结构破坏严重,墙体开裂,并有明显的滑移、错位,部分墙体外闪,个别墙板塌落
毁坏	V	外墙体近乎全部倒塌,承重结构部分倒塌,需拆除重建

(1) 地震特性:地震时的地面运动动力特别是地面运动水平加速度峰值是造成建筑物破坏的主要因素^[3],本文选取地面加速度峰值作为地震特性指标。

(2) 场地与地基条件:场地与地基条件越差,房屋的震害就越严重。

(3) 建筑物层数与层高:房屋的层数与层高不仅影响房屋的空间刚度,而且使房屋受到的地震力的大小亦大为不同,层数越多,层高越高,其房屋的震害就越重。

(4) 含砖墙率:含砖墙率越大房屋的空间刚度越大,抗剪能力越强,因此震害就越轻。

(5) 其它抗震性能:包括结构的合理性、房屋的整体性、房屋质量、使用年限等因素。横墙承重或纵横墙承重的房屋比纵墙承重的房屋抗震能力要强;圈梁的道数越多震害越轻;砂浆标号越高,震害越轻。

本文在提出特征向量建立预测模型时,除考虑各因子对结构的抗震性能的影响外,还需考虑参数的易获得性和独立性。

2.2 数据离散化

RS理论只能处理离散属性,如果某些条件属性或决策属性的值域为连续值,则必须对属性值进行离散化处理;即便某些属性值是离散数据,有时也需要通过对离散值进行合并得到更高抽象层次的离散值。数据离散化本质上可归结为利用选取的断点来对条件属性构成的空间进行划分的问题,即把条件属性划分成有限个区域,使得每个区域中的对象的决策值相同。数据离散可根据实际情况,选择等距离算法、等频率算法、Naive Scaler算法等^[11]。

2.3 属性约简、粗糙依赖度及重要度计算

一般情况下,知识库中的知识并不是同等重要的,可能存在冗余,尤其当知识库数据系随机采集时其冗余性更普遍。冗余知识的存在一方面增大工作量,另一方面会干扰人们做出正确而简明的判断。知识约简就是在保持分类能力不变的情况下,删除其中不相关或不重要的知识,得到简洁的知识表达。知识约简是RS理论的核心内容。可利用粗糙依赖度衡量决策变量与属性之间的依赖程度,利用重要度衡量某属性对决策的影响程度。

定义2(属性约简) 若 $P, Q \subseteq R$,若 $ind(P) \subseteq ind(Q)$,则称知识 Q 依赖知识 P ,也就是说知识 Q 可由知识 P 完全推出,记为 $P \rightarrow Q$ 。当 $a \in P$,且 $pos_P(Q) = pos_{(P-\{a\})}(Q)$ 时,称 a 为 P 中 Q 不必要的

(冗余的),则 a 可以被约简,否则 a 为 P 中 Q 必要的。

定义3(粗糙依赖度及重要度) 知识的依赖度 $\gamma = \gamma_P(Q) = |pos_P(Q)| / |U|$,衡量知识 Q 对知识 P 的依赖程度($0 \leq \gamma \leq 1$)。当 $\gamma = 1$ 时,知识 Q 完全依赖 P ;通过 γ 值来确定决策属性与条件属性之间的依赖关系。条件属性 C 的子集 C' 关于决策属性 D 的重要度定义为 $\sigma_{CD}(C') = \gamma_C(D) - \gamma_{C-C'}(D)$ 。属性的重要度越大,该属性对知识分类的影响就越大。

2.4 粗集-神经网络的震害预测模型

粗集-神经网络的震害预测模型由两部分构成,将粗糙网络作为前置系统,神经网络作为后置的信息识别系统。首先用粗糙集方法对信息进行预处理,然后根据预处理所抽取的关键成分作为神经网络的输入,训练神经网络,建立输入与输出的非线性关系。由于通过粗糙集方法去掉冗余信息后,减少了信息表达的特征数量,简化网络结构,缩短训练时间。而神经网络具有较强的容错及抗干扰能力,可以很好地消除粗糙集对噪音敏感的缺陷。因此粗糙神经网络综合了粗糙集和神经网络的优点,可以很好解决非线性的预测问题。

粗糙-神经网络结构如图1所示,本文选择震害影响因子:房屋层数 c_1 、房屋质量 c_2 、砂浆标号 c_3 、砖墙面积率 c_4 、房屋整体性 c_5 、场地条件 c_6 、地震峰值加速度 c_7 为模型的输入,将给定地震下可能的建筑物的震害程度为模型的输出。建模过程分为四步:

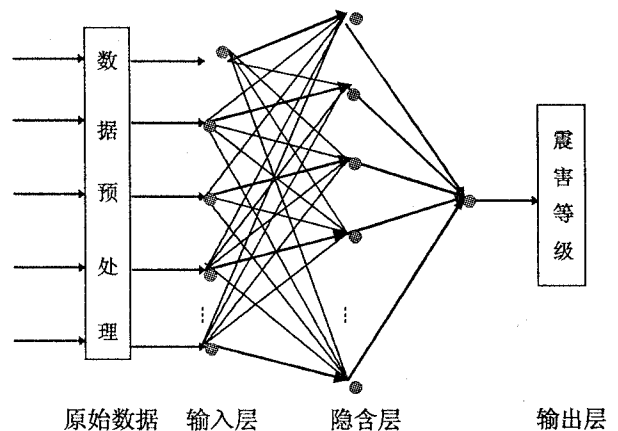


图1 震害预测的粗糙神经网络结构示意图

Fig.1 Sketch of seismic damage prediction to buildings based on the rough set and neural network.

(1) 连续型数据离散化处理:本文采用部分采用等距离算法,部分因子采用专家经验法进行离散处理。

(2) 属性的约简:约简的一般方法为:①在决策表中将信息相同(即具有相同描述)的对象及其信息删除,只留其中一个得到压缩后的信息表,即删除多余事例;②删除多余的属性;③求出最小约简。

(3) 神经网络模型训练:把经过简约后的数据信息表输入到BP神经网络中进行训练。

(4) 建筑物可能震害程度的预测。

3 实例及分析

笔者收集了1978年唐山大地震、1966年云南东川地震、1988年澜沧-耿马地震、1995年武定地震和1997年丽江地震等有关多层砖房震害资料^[1,3,6],选择其中的93个样本训练基于粗糙神经网络多层砖房震害模型。

参考文献[2,5]选取房屋层数、施工质量、砂浆标号、砖墙面积率、房屋整体性、场地条件、地震峰值加速度共7个初始影响因子作为模型输入。对于定性指标参考专家经验赋值,如房屋质量:施工质量优为10、中取8、差取6,同时考虑房屋折旧,每5年减0.5。房屋的整体性:楼盖和屋盖,现浇取5、预制4、木屋盖3,有圈梁加1、有构造柱再加1,有地下室、筏基加1,房屋开裂减1。场地条件:I类场地取1、II类取2、III类取3、IV类取4。震害作为网络最终输出,把震害程度划分为五级:基本完好、轻微破坏、中等破坏、严重破坏、毁坏,所收集的样本资料如表3(限于篇幅仅列出部分样本)。

表3 多层砖房震害样本的输入与输出

序号	房屋层数 c_1	房屋质量 c_2	砂浆标号 c_3	砖墙面积率 c_4	房屋整体性 c_5	场地条件 c_6	峰值加速度 c_7	震害类型 d
1	2	10	25	8.2	8	3	0.20	III
2	2	6	10	10.0	8	4	0.25	IV
3	2	8	20	9.0	12.0	1	0.10	I
4	2	10	25	9.7	9.5	2	0.15	II
5	4	10	25	7.6	7.5	1	0.20	III
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
93	5	8	10	9.0	9.0	1	0.15	II

震害预测信息系统 $S = (U, R, V, f)$ 中,将研究区域内的全部建筑物组成的集合作为论域 U ,震害影响因子作为条件属性 $C = \{c_1, c_2, c_3, c_4, c_5, c_6, c_7, c_8\}$,给定地震加速度峰值下建筑物可能的破坏程度为决策属性 $D = \{d\}$,属性集 $R = C \cup D, C \cap D = \phi$ 。决策表的表头为各属性,其每一行表示论域中的一个事例或称一条决策规则,每一列表示属性及属性值。实例的属性值按表4进行离散化后,形成初始决策表5。

利用初始决策表 U_0 计算可得决策属性对条件属性的依赖度 $\gamma_{C_0}(D_0) = 1$,证明决策属性 d 完全依赖于条件属性 $(c_1, c_2, c_3, c_4, c_5, c_6, c_7, c_8)$,说明了指标选取及其离散化是合理的。进一步分析得 $U_0/(C_0 - \{c_1\}) \neq U_0/(C_0), U_0/(C_0 - \{c_2\}) = U_0/(C_0), U_0/(C_0 - \{c_3\}) = U_0/(C_0), U_0/(C_0 - \{c_4\}) \neq U_0/(C_0), U_0/(C_0 - \{c_5\}) \neq U_0/(C_0), U_0/(C_0 - \{c_6\}) \neq U_0/(C_0), U_0/(C_0 - \{c_7\}) \neq U_0/(C_0)$,表明条件属性 c_2 和 c_3 是不必要的(属冗余信息),约去 c_2 和 c_3 不会影响分类能力。将 c_2 和 c_3 从决策表中删除后形成新决策表 $U_1, D_1 = \{d\}, C_1 = \{c_1, c_4, c_5, c_6, c_7\}$,计算得依赖度 $\gamma_{C_1}(D_1) = 1, \gamma_{(C_1 - \{c_1\})}(D_1) = 0.91, \gamma_{(C_1 - \{c_4\})}(D_1) = 0.60, \gamma_{(C_1 - \{c_5\})}(D_1) = 0.69, \gamma_{(C_1 - \{c_6\})}(D_1) = 0.73, \gamma_{(C_1 - \{c_7\})}(D_1) = 0.49$; 据计算可得各条件属性(即震害影响因子)对决策属性(建筑物震害程度)的重要度: $\sigma_{C_1, D_1}(c_1) = 0.09, \sigma_{C_1, D_1}(c_4) = 0.40, \sigma_{C_1, D_1}(c_5) = 0.31, \sigma_{C_1, D_1}(c_6) = 0.27, \sigma_{C_1, D_1}(c_7) = 0.51$ 。各因子的重要度排序为: $c_7 > c_4 > c_5 > c_6 > c_1$ 。

表4 属性值及其离散化

V	c_1	c_2	c_3	c_4	c_5	c_6	c_7	d
0	[2,3]	10	[30,35]	[12.0,15.0]	[10,12]	1	[0.1,0.15]	I
1	[3,4]	8	[20,25]	[9.0,12.0]	[8,10]	2	[0.15,0.20]	II
2	[4,5]	6	[5,15]	[6.0,9.0]	[6,8]	3	[0.20,0.25]	III
3		4		[3.0,6.0]		4	[0.25,0.30]	IV
4							[0.30,0.35]	V

表5 初始决策表

U_0	c_1	c_2	c_3	c_4	c_5	c_6	c_7	d
u_1	0	0	1	3	1	2	2	2
u_2	0	3	2	1	1	1	3	3
u_3	1	0	1	2	0	0	2	0
u_4	0	0	1	1	1	1	1	1
u_5	2	0	1	2	2	0	2	2
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
u_{93}	2	1	1	2	1	0	1	1

本文对决策表进行约简后得到51条规则。用试算法确定最佳隐层个数,得最优神经网络结构为5-11-1。利用Matlab神经网络工具箱,其中选择隐含层传递函数为“Logsig”,输出层传递函数为“purelin”,训练网络调用的函数为train(net, x, y),训练过程显示频率设为100,最大训练步数设置为10000,误差指标设置为0.01。当训练次数为795次时,训练误差已低于 10^{-5} ,表明用粗糙集理论进行条件属性和对象约简,去掉冗余的属性和对象,虽初始样本多达93个,但通过粗糙集理论分析仅保留了51条分类规则作为最终模型的输入,基于粗糙集

理论框架下构建的神经网络结构大大简化,提高了计算效率和计算精度。

为了验证模型的性能,另选 10 个样本进行测试,计算结果如表 6。其中有 9 个样本的计算结果与实测结果完全吻合,8 号样本的预测结果有偏差,可能是由于建模所采用的原始样本(库知识)中缺乏与 8 号样本相近的样本。这也说明在选取样本建模时要特别注意原始样本的代表性。

表 6 10 个样本的预测结果

序 号	房屋 层数 c_1	房屋 质量 c_2	砂浆标 号 c_3	砖墙面 积率 c_4	房屋整 体性 c_5	场地条 件 c_6	峰值加 速度 c_7	目标输 出 d	计算 结果 d'
1	3	8	20	6.7	8.0	4	0.25	IV	IV
2	2	10	25	6.5	7.5	2	0.10	II	II
3	4	6	10	4.6	7.5	5	0.15	IV	IV
4	4	8	25	8.5	9.5	3	0.30	IV	IV
5	2	9	25	10.3	7.0	2	0.25	III	III
6	3	10	10	8.0	6.0	3	0.30	IV	IV
7	2	8	10	4.5	9.0	1	0.10	II	II
8*	4	6	10	6.5	6.5	4	0.30	V	IV
9	2	9	25	9.3	8.5	3	0.20	II	II
10	5	10	10	11.5	8.0	2	0.25	I	I

4 结论

震害预测是城市防震减灾的基础工作,准确的建筑物震害的震前预测和震后评价可提高城市防震减灾工作的科学性。基于粗糙集的神经网络建筑物震害预测模型综合了粗糙集理论在知识获取方面的能力和神经网络在数值逼近上的优势。震害预测工作涉及到前期大量数据的现场收集与调研以及后期的数据处理,将粗糙集理论的属性约简用于震害预测工作中,这将大大减少数据收集与处理的工作量。

通过粗糙集智能数据分析,消除初始决策表中的冗余信息和噪声数据的干扰,精简规则集,简化神经网络的结构,缩短训练时间,提高泛化能力。

实例研究表明,对多层砖房震害预测与实际震害基本吻合,提高了建筑物震害预测的可靠性。该预测模型会随着新样本数据的加入而增加新的判据,具有学习、修正、更新分类规则的功能,使预测系统逐步完善。

[参考文献]

- [1] 刘恢先.唐山大地震城害[M].北京:地震出版社,1986.
- [2] 尹之潜.地震灾害及损失预测方法[M].北京:地震出版社,1996.
- [3] 汤皓,陈国兴,李方明.基于BP神经网络模型的多层砖房震害预测方法[J].地震工程与工程振动,2006,(4):141-146.
- [4] 赵艳林,杨绿峰.多层砖房震害的灰色预测[J].世界地震工程,2002,18(1):56-61.
- [5] 阎维明,乔亚玲,何浩祥.粗糙集理论在震害预测中的应用[J].自然灾害学报,2006,15(3):147-151.
- [6] 白良.澜沧-耿马地震多层砖房的震害及机理简析[J].云南工学院学报,1989,4(3):35-43.
- [7] 石玉成,马尔曼,陈永明,等.2003年民乐-山丹6.1、5.8级地震房屋震害分析[J].西北地震学报,2005,27(3):260-266.
- [8] 马尔曼,石玉成,高晓明,等.2003年民乐-山丹6.1、5.8级地震灾害损失评估[J].西北地震学报,2007,29(1):24-27.
- [9] 马尔曼,陈永明,赵广奎,等.2002年甘肃玉门5.9级地震灾害损失评估[J].西北地震学报,2004,26(2):162-167.
- [10] 许东,吴铮.基于MATLAB的系统分析与设计-神经网络[M].西安:西安电子科技大学出版社,2002.
- [11] 张文修,吴伟志.粗糙集理论与方法[M].北京:科学出版社,2001.