·信息安全技术·

**熵权神经网络的信息系统安全评估模型**

顾兆军1,2 辛倩1,2

（1．中国民航大学，信息安全测评中心，天津 300300；2．中国民航大学 计算机科学与技术学院，天津 300300）

**摘 要：**信息系统安全影响因素众多，且因素之间互相影响，呈现出复杂的非线性关系。传统的数学评估模型无法解决因素之间的关联性以及传统BP神经网络易出现过拟合现象，从而导致评估结果精度较低。为提高信息系统的安全评估精度，提出了一种基于熵权法的BP神经网络评估模型。首先采用熵权法对安全指标体系进行处理，提取重要指标，剔除对评估结果几乎影响的指标，然后将重要指标作为BP神经网络的输入向量进行训练，降低BP神经网络的输入维数，提高BP神经网络对知识的学习能力，最后利用实例验证此模型的性能。结果表明，与其他同类型评估模型相比，熵权-BP模型训练速度快，评估精度高，可以为信息系统管理员提供科学的参考依据。

**关键词：**信息系统安全；熵权法；神经网络；评估模型；非线性

**中图法分类号：** TP393.08 **文献标识码：**A **DOI:**

Information System Security Evaluation Based on Entropy Method and Neural Network

Zhao Jun Gu1,2 Qian Xin1,2

(1．Information Security Evaluation Center, Civil Aviation University of China, Tianjin 300300; 2. college of Computer Science and Technology, Civil Aviation University of China, Tianjin 300300）

**Abstract:** The security of information system is affected by many factors, and the factors affect each other, showing a complex nonlinear relationship.The traditional mathematical evaluation model can not solve the correlation between the factors, and the traditional BP neural network prone to overfitting phenomenon, resulting in lower accuracy of the evaluation results. In order to improve the accuracy of information system security assessment, a BP neural network evaluation model based on entropy weight method is proposed in this paper. Firstly, the entropy method is used to deal with the safety index system to extract the important indexes and eliminate the indicators that have little impact on the evaluation results, and secondly the important index is trained as the input vector of BP neural network to reduce the input dimension of BP neural network, improve the learning ability of BP neural network, the performance of this model is verified in the end. The results show that compared with other similar types of evaluation models, the evaluation accuracy of entropy weight -BP model is higher and speed is faster. The method can provide a scientific reference for information system administrators.

**Key words:** information system; entropy weight method; neural network evaluation model; nonlinearity

# 引言

随着各行各业对信息系统的依赖程度日益增加，一些重要的信息系统已经成为国家重点保障的

关键信息基础设施。然而国内外不断地出现严重的信息系统安全事件，如浙江温州有线电视网络系统被黑客攻击，乌克兰电力信息系统遭受恶意代码攻击，以及最近爆发的全球大规模蠕虫勒索软件感染事件等，这些安全事件将会给企业甚至国家带来无法估计的损失和影响。因此，信息系统的安全状况成为人们当前关注的热点问题，如何准确地评估信息系统的安全状况对保障信息系统安全具有重要的研究价值和意义。

信息系统安全评估方法主要分为定性评估方法、定量评估方法和综合评估方法。由于影响信息系统的安全因素难以量化，大多数学者采用层次分析法、熵权法、模糊理论、故障树等[1,4]定性与定量相结合的综合评估方法建立信息系统安全评估模型，但这些方法计算复杂、评估结果受主观因素影响较大，且无法解决非线性问题。为了避免这些缺陷，国内外学者采用DS证据理论[5]、支持向量机[6]、神经网络[7]、贝叶斯网络[8]等人工智能算法构建信息系统安全评估模型，取得了较好的评估结果。de Gusmão A P H [9]将事件树与模糊决策理论相结合，通过专家鉴定，确定场景和事件，模糊评价和排序这一系列步骤对信息安全进行分析，但在确定事件发生时没有统一的方法，导致不同方法设置事件发生的概率不同；马丽仪[10]等人采用模糊神经网络评估信息系统的安全风险，将风险因子进行模糊化后输入到神经网络训练，提高了神经网络对模糊性的识别能力，但隶属度矩阵和权重均采用专家评判，客观性较差，在一定程度上影响整个模型的实用性；赵保华[11]采用层次分析法和神经网络构建信息系统风险评估模型，层次分析法确定评估指标重要性时完全依赖评估者的主观判断，可能会误剔重要指标，而且评估指标较多容易出现残缺判断矩阵；赵刚[12]等人使用熵权法计算指标权重，使模糊合成中的权重更加客观，但是当信息系统有较多互相关联的安全因素时，缺乏学习和自适应能力。

为了进一步提高评估的准确性，针对当前信息系统安全影响因素较多、且各因素之间相互影响等特点，本文提出了熵权法优化BP神经网络（Back-Propagation Neural Network）的信息系统安全评估模型（熵权-BP），并采用具体的实例验证模型的有效性。

# 信息系统安全评估模型

信息系统安全的安全因素众多、且因素之间相互关联，这些影响因素与信息系统安全评估之间存在着非常复杂的非线性关系，传统的数学某型无法处理这种非线性的映射。信息系统安全评估安全评估模型可以表示为：

 (1)

式中，*Y*是信息系统安全评估结果，*xi*表示影响信息系统安全的因素。*f* ( )为非线性函数，所以信息系统安全评估结果的准确性与安全因素的选取以及非线性函数*f* ( )直接相关。本文选择熵权法对评估指标进行重构，并采用具有高度非线性能力的BP神经网络建立信息系统安全评估模型，熵权-BP模型流程如图1所示。



图1 熵权-BP安全评估模型流程图

# 2 熵权-BP神经网络信息系统安全评估模型

## 2.1 构建评估指标体系

评估信息系统的安全状况首先是构建评估指标，即信息系统的安全影响因素。安全影响因素的选取是

否科学合理，直接影响评估结果的真实性。本文根据《信息安全技术 信息系统安全等级保护测评要求》[13]，从技术和管理两大层面构建了层次结构的信息系统安全评估指标体系，如图2所示。其中技术层面包括物理安全、网络安全和运行安全，管理层面包括制度、机构、人员、建设和运维安全管理。



图2 信息系统安全评估指标体系

## 2.2 熵权法筛选重要指标

由于影响信息系统安全的因素众多，必然存在对评估结果无影响的因素，为避免无效信息对BP神经网络学习过程的干扰，应提取重要的指标作为BP神经网络的输入向量。在对信息系统安全评估时，通常用权重来表示指标的重要性。熵权法属于客观赋权法，通过实际数据利用严谨的数学理论得到较为客观的指标权重，避免了主观方法确定权重的严重缺陷，使指标权重更具有科学性和说服力。

信息论中的熵值反映了信息的无序化程度，可用来度量数据所提供信息量的大小。当某项指标在不同系统的变异度越大时，熵值越小，说明该指标传递的有效信息量就越大，在综合评估中起的作用就越大，则该指标的权重也应该越大；反之某项指标的变异度越小，熵值越大，说明该指标传递有效信息量就越小，在综合评估中起的作用就越小，该指标的权重也应越小。因此，本文根据同一评估指标在不同信息系统的变异度，利用信息熵计算出各评估指标的权重。

2.2.1构建评估指标矩阵

已完成评估的*m*个具有代表性的信息系统与*n*个评

估指标形成信息系统评估指标矩阵***R***=(*rij*)*nm*：

 (2)

其中*rij*(*i*=1,2,..,*n*;*j*=1,2,…,*m*)为第*j*个信息系统的第*i*个指标的实际评估结果。

2.2.2 计算评估指标的权重

（1）根据评估指标矩阵***R***，计算第*i*个评估指标下第*j*个信息系统的评估指标值的比重*pi*j：

 (3)

（2）计算第*i*个评估指标的熵权*ei*：

  (4)

（3）用ln(*m*)对各评估指标的熵权进行归一化处理：

 (5)

（4）引入差异系数1*-ei* 来正向度量评估指标*i*的熵权，得到评估指标*i*的权重：

 (6)

## 2.3 熵权-BP神经网络结构的构建

BP神经网络即误差反向传播神经网络，具有较强的自学习和自适应能力，目前被广泛用在综合评估领域。由于信息系统安全因素较多，且因素之间相互影响，为了避免BP神经网络在学习过程中陷入过拟合，采用熵权法提取出重要的评估指标作为BP神经网络的输入，消除无效指标的干扰，降低BP神经网络的输入维数，提高评估指标体系的质量，从而提高BP神经网络的学习能力。熵权-BP神经网络结构如，图3所示。



图3 熵权-BP神经网络结构

图中*u*表示指标体系的所有评估指标，*x*表示重要的评估指标，*y*表示隐含层神经元的输出，*z*表示输出层神经元的网络输出值。*wij*和*wj*分别为传递层到隐含层与隐含层到输出层之间的连接权值。*l、n*和*m*分别比为输入层、传递层和隐含层神经元的个数。

参照上图，熵权-BP神经网络建模过程：

**Step1：**第1层为输入层，其神经元的个数为信息系统安全评估指标体系中的19个评估指标；

**Step2：**第2层为指标筛选层，通过构造评估指标矩阵利用信息熵计算第1层评估指标的权重，并筛选出重要评估指标；

**Step3：**第3层为传递层，将第2层选取出重要的评估指标传递到BP神经网络中；

**Step4：**第4层为隐含层，主要处理评估指标到评估结果间的非线性映射，隐含层神经元个数的选取是BP神经

网络学习成功的关键，借助经验公式：

 (7)

式中：*n*表示传递层神经元个数，*β*∈[1,10]的常数。

**Step5：**第5层是输出层，根据实际需要只设定一个神经元，代表信息系统安全评估结果。

此时，熵权-BP神经网络结构的构建已完成，接下里就是第3层到第4层之间与第4层到第5层之间权值的学习过程。

## 2.4 BP神经网络学习算法

如图4所示，传递层神经元用*i*表示，隐含层神经元用*j*表示。给定*h*组样本的输入和输出分别为：***X***=(***X***1,***X***2,…,***X***3)*T*=(***Xpi***)*h*×*n*、***z’p***=( *z’*1, *z’*2,…, *z’h*) (*p*=1,2,…,*h*)。

BP神经网络学习过程：

（1）随机初始化BP神经网络所有的权值和阈值；

（2）BP神经网络正向传递信号，传递层到隐含层神经元样本*p*的计算公式：

 (8)

隐含层到输出层神经元样本*p*的计算公式：

 (9)

神经元作用函数为：

 (10)

式中，*θj*表示隐含层神经元*j*(*j=*1,2,…,*m*)的阈值，*θ*表示输出层神经元的阈值；

（3）使用α-η联合算法反向修正权值和阈值

传递层到隐含层以及隐含层到输出层之间的连接权值修正：

 (11)

隐含层和输出层的阈值修正：

 (12)

其中*δ* 和*δk* 为误差项，输出层和隐含层神经元误差项：

 (13)

式中，*f’*为对*f*求导，*η*∈(0,1)为学习速率，加快网络的收敛速度；*α*∈(0,1)为动量因子，避免BP网络陷入局部最优。

（4）*h*组样本拟合误差函数为：

 (14)

 (15)

 (16)

式中，*ε*为误差精度，满足(0≤*ε*<1)，当网络误差达到精度要求或者达到学习的最大次数*T*(*t*≤*T*)，BP网络学习过程结束，否则转向步骤(2)，继续学习直到满足条件为止。

# 3 实例研究

## 3.1 数据来源

本文收集民航已完成安全测评的30个信息系统作为样本数据，如表1。选择前22个作为熵权-BP模型的训练样本，其余8个作为测试样本，用于检验熵权-BP模型的安全评估效果。

表1 信息系统安全评估数据

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 样本编号 | *u1* | *u2* | *u3* | *u4* |  | *u19* | *z* |
| 1 | 68.4 | 59.43 | 68.43 | 62.7 |  | 80.97 | 0.754 |
| 2 | 75.4 | 65.2 | 75.67 | 82.84 |  | 84.62 | 0.813 |
| 3 | 61.64 | 62.59 | 75.11 | 65.8 |  | 87.37 | 0.676 |
| 4 | 69.28 | 59.66 | 72.35 | 60.27 |  | 76.87 | 0.755 |
| 5 | 70.16 | 73.53 | 75.36 | 72.66 |  | 68.25 | 0.716 |
| 6 | 91.93 | 67.29 | 63.4 | 82.24 |  | 78.72 | 0.832 |
|  |  |  |  |  |  |  |  |
| 30 | 63.42 | 65.9 | 61.25 | 71.23 |  | 63.86 | 0.677 |

## 3.2 筛选重要评估指标

（1）使用已完成安全测评的30组评估指标数据形成一个评估指标矩阵***R***,即



（2）通过2.3节的计算步骤，得出各评估指标的权重*W*=(*w1*,*w2*,…,*w14*)：

***W***=(0.17,0.008,0.006,0.24,0.02,0.03, 0.03, 0.039,0.04,0.1, 0.17, 0.09, 0.01, 0.01,0.02,0.035,0.05, 0.08 , 0.06)

对评估指标权重进行排序，经验证改变权重小于0.05指标的取值对整个评估结果几乎没有影响，可以作为数据噪声将其剔除，剩下8个重要指标 ：

***W’***=(0.24,0.17,0.17,0.1,0.09,0.08,0.06,0.05)，其分别代表网络设备安全、物理设备安全、应用软件安全、终端设备安全、数据库安全、人员安全管理、运维安全管理和建设安全管理8个重要指标。

## 3.3 熵权-BP神经网络模型训练

将重要指标数据采用最大最小法归一化处理后输入到BP模型中进行学习训练，BP神经网络的学习率*η*为0.5，动量因子*α*为0.85，最大学习次数*T*为6000，拟合误差*ε*为0.001，根据经验公式得到隐含层神经元个数为6个。当训练400次时，拟合误差已经达到0.001，此时熵权-BP模型误差已达到预设要求。为了比较熵权-BP模型与BP模型的学习能力，同时对两个模型进行训练，其训练过程如图4所示。

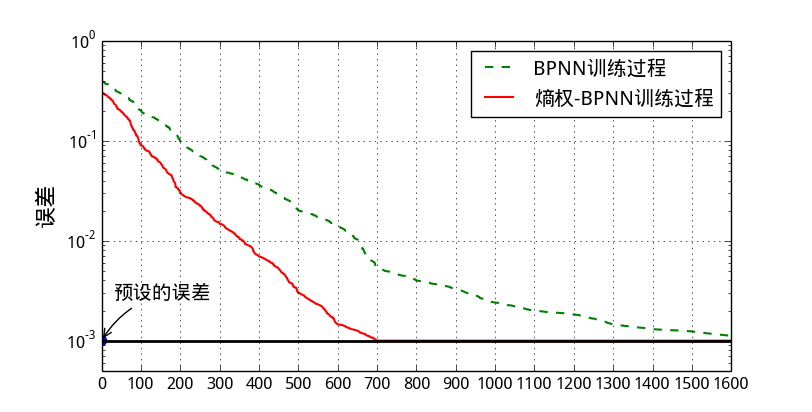


图4 BP神经网络训练过程

由图4可知，在相同的误差学习目标下，基于原始指标数据的BP神经网络模型的训练次数明显高于以熵权重构后的数据为输入的BP模型的训练次数。

训练结果表明，熵权-BP模型通过对8个重要的评估指标数据进行训练，既保留了重要指标的特征，又减少了无用指标对BP神经网络学习的干扰，可以提高BP神经网络的训练速度。

## 3.4 熵权-BP神经网络模型检验

为了检验熵权-BP模型的评估结果，将测试样本分别输入到BPNN、AHP-BP和熵权-BP模型中进行安全评估。评估结果和评估精度分析结果分别为如图5和表2。

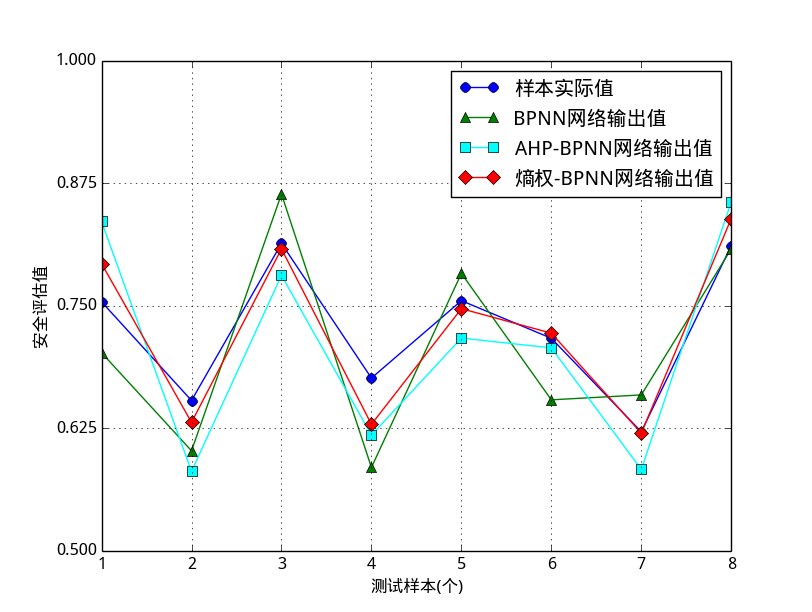


图5 各模型的网络输出值和期望值之间的对比

表2 各模型的信息系统安全评估精度对比

|  |  |
| --- | --- |
| 模型 | 评估精度/% |
| BPNN | 86.72 |
| AHP-BP | 91.39 |
| 熵权-BP | 96.41 |

从图5与表2可知，组合的熵权-BP和AHP-BP模型的评估精度高于单一BPNN模型的评估精度，因为这两种组合模型都对评估指标进行了重要性筛选，使得BP神经网络在反向传播过程中更加逼近真实权值；熵权-BP模型的评估精度高于AHP-BP模型的评估精度，因为熵权-BP模型通过样本数据计算得到指标的权重，更加客观、真实的反映评估指标的权重，而AHP-BP模型通过专家构造判断矩阵计算指标权重，指标的权重受主观因素影响较大，可能会错误删除重要的指标。总之，熵权-BP模型优于其他两种模型，利用其进行信息系统安全评估是可行的。

# 4 结束语

信息系统安全影响因素众多、且各因素之间相互影响，传统的BP神经网络存在学习速度慢、过拟合不足，为提高信息系统安全评估的精度，提出一种熵权神经网络的信息系统安全评估模型。熵权法可以对BP神经网络模型的输入数据进行重构，在保留影响因素最大信息量的前提下，降低BP神经网络的输入维数，简化了BP网络结构，同时具有高度非线性处理能力的BP神经网络，能有效解决各因素之间的关联性，弥补传统评估方法无法处理非线性关系的不足。但在构建评估指标体系处还存在缺点，本文只是从等保测评的要求中选取了具有代表性的几大类安全影响因素，而每类因素还包括许多具体的因子，不同安全等级的信息系统每个因素的具体因子可能会有些许不同。后续需要结合实际情况对各类因素进行细分，使评估结果能够为系统管理员提供更有效的参考依据。

# 参考文献

1. Zhang L, Peng J, Du Y, et al. Information security risk assessment survey[J]. Journal of Tsinghua University Science and Technology, 2012, 52(10): 1364-1369. (in Chinese) [张利, 彭建芬, 杜宇鸽, 等. 信息安全风险评估的综合评估方法综述[J]. 清华大学学报: 自然科学版, 2012, 52(10): 1364-1369.]
2. Gu Z, Han Y. Quantifiable information system security level evaluation[J]. Computer Engineering and Design, 2016, 37(7): 1729-1733. (in Chinese) [顾兆军, 韩迎亚. 可量化的信息系统安全水平评估[J]. 计算机工程与设计, 2016, 37(7): 1729-1733.]
3. Rosqvist T, Molarius R, Virta H, et al. Event tree analysis for flood protection—An exploratory study in Finland[J]. Reliability Engineering & System Safety, 2013, 112: 1-7.
4. Lin Y. Research on Information System Security Risk Assessment Modeling and Application[D]. Dalian: Dalian Maritime University, 2013.(in Chinese) [林洋. 信息系统安全风险评估模型及其应用研究[D]. 大连: 大连海事大学, 2013.]
5. Shi B, Xie X Q. Research on network security situation forecast method based on D-S evidence theory[J]. Computer Engineering and Design, 2013, 34(3): 821-825.(in Chinese) [石波, 谢小权. 基于 DS 证据理论的网络安全态势预测方法研究[J]. 计算机工程与设计, 2013, 34(3): 821-825.]
6. Wu L. Risk Analysis of the Information System by Using Factor Analysis and Support Vector Machine[J]. Microelectronics and Computer, 2016, 33(2): 144-148.(in Chinese) [伍浏阳. 因子分析和支持向量机的信息系统风险评估[J]. 微电子学与计算机, 2016, 33(2): 144-148.]
7. Gao N, Gao L, He Y, et al. Dynamic Security Ｒisk Assessment Model Based on Bayesian Attack Graph[J]. Journal of Sichuan University: Engineering Science Edition, 2016, 48(1): 111-118(in Chinese) [高妮, 高岭, 贺毅岳, 等. 基于贝叶斯攻击图的动态安全风险评估模型[J]. 四川大学学报: 工程科学版, 2016, 48(1): 111-118.]
8. Yu F, Xiaoping W, Jia-Sheng W. An approach to information systems security risk assessment based on fuzzy-combinatorial neural network[J]. Journal of Naval University of Engineering, 2010, 1: 18-23. (in Chinese) [付钰, 吴晓平, 王甲生. 基于模糊-组合神经网络的信息系统安全风险评估[J]. 海军工程大学学报, 2010 (1): 18-23.]
9. de Gusmão A P H, e Silva L C, Silva M M, et al. Information security risk analysis model using fuzzy decision theory[J]. International Journal of Information Management, 2016, 36(1): 25-34.
10. Ma L Y, Zhang L F, Yang Y, et al. Evaluation of information system security risk based on fuzzy neural network method[J]. Zhongguo Anquan Kexue Xuebao, 2012, 22(5): 164-169. (in Chinese) . [马丽仪, 张露凡, 杨宜, 等. 基于模糊神经网络方法的信息系统安全风险评估研究[J]. 中国安全科学学报, 2012, 22(5): 164-169. ]
11. Zhao B. Risk Evaluation of Information System Security Based on Neural Network and Analytic Hierarchy Process[J]. Microelectronics and Computer, 2015, 32(10): 163-166. (in Chinese) [赵保华. 层次分析法和神经网络的信息系统风险评估[J]. 微电子学与计算机, 2015, 32(10): 163-166.]
12. Zhao G, Liu H. Practical risk assessment based on multiple fuzzy comprehensive evaluations and entropy weighting[J]. Journal of Tsinghua University Science and Technology, 2012, 52(10): 1382-1387. (in Chinese). [赵刚, 刘换. 基于多层次模糊综合评判及熵权理论的实用风险评估[J]. 清华大学学报: 自然科学版, 2012, 52(10): 1382-1387.]
13. 2012 GBT. Information security technology - Information system security level protection evaluation requirements[S][D]. (in Chinese) [2012 G B T. 信息安全技术-信息系统安全等级保护测评要求 [S][D]]

# 作者简介：

顾兆军(1966-)，男，山东省蓬莱市人，博士，教授，研究方向为网络与信息安全、民航信息系统；

通讯作者：辛倩(1992-)，女，山西省大同市人，硕士研究生，研究方向网络与信息安全、民航信息系统。

# 联系方式：

联系人：辛倩  
联系电话：13516169009

E-mail:qxinlu1992@126.com  
邮编：300300  
地址： 天津市东丽区津北公路2898号中国民航大学北院48号信箱