(19)中华人民共和国国家知识产权局

(12)发明专利

(10)授权公告号 CN 103581188 B
(45)授权公告日 2016.08.03

(21)申请号 201310544315.8

(22)申请日 2013.11.05

(73)专利权人 中国科学院计算技术研究所
地址 100190 北京市海淀区中关村科学院南路6号

(72)发明人 金舒原 庞依 张亚星

(74)专利代理机构 北京律诚同业知识产权代理有限公司 11006
代理人 祁建国 梁挥

(51)Int.Cl.
H04L 29/06(2006.01)
G06N 3/02(2006.01)
G06N 3/12(2006.01)

(56)对比文件
JP 2007049591 A,2007.02.22,
CN 102799627 A,2012.11.28,
CN 103295081 A,2013.09.11,
CN 102340485 A,2012.02.01,
李凯,曹阳.基于ARIMA模型的网络安全威胁态势预测方法.《计算机应用研究》.2012,全文.
李志东,杨武.基于场景平移的网络安全态势预测.《高技术通讯》.2012,全文.
王慧强,赖积保.网络态势感知系统研究综述.《计算机科学》.2006,全文.

审查员 何思佳

权利要求书5页 说明书15页 附图3页

(54)发明名称

一种网络安全态势预测方法及系统

(57)摘要

本发明涉及一种网络安全态势预测方法及系统,方法包括:将获得的网络安全态势值序列集合作为训练数据;对反向传播神经网络结构进行初始化,包括设定输入层神经元个数M和输出层神经元个数N;对该训练数据进行实数编码,并找到最具适应度训练数据;将该最具适应度训练数据中的该输入层神经元个数M所对应的安全态势值作为输入值,该输出层神经元个数N所对应的安全态势值作为期望输出值,训练该反向传播神经网络,并建立网络安全态势的预测模型;将该输入层神经元个数M所对应的安全态势值作为输入值,根据该预测模型预测该输出层神经元个数N所对应的网络安全态势值。该方法能提高网络安全态势预测的收敛速度,降低训练时间和预测误差。

CN 103581188 B

1.一种网络安全态势预测方法,其特征在于,包括以下步骤:

步骤1,将通过采集并融合局域网内资产、流量、入侵检测系统警报、漏洞数据而计算出的多个网络安全态势值作为训练数据;

步骤2,对反向传播神经网络结构进行初始化,包括设定输入层神经元个数M、隐藏层神经元个数L和输出层神经元个数N;

步骤3,对该训练数据进行长度为K的实数编码,其中,

K=权值个数+偏倚个数=(M*L+L*N)+(L+N),

则每个编码后的该训练数据包含该反向传播神经网络的权值和偏倚信息,由适应度函数计算编码后的该训练数据的适应度值所确定的概率,找到最具适应度训练数据;

步骤4,将该最具适应度训练数据中的该输入层神经元个数M所对应的安全态势值作为输入值,该输出层神经元个数N所对应的安全态势值作为期望输出值,根据向前传递该输入值、向后传播该期望输出值来训练该反向传播神经网络,从而建立网络安全态势的预测模型;

步骤5,将该输入层神经元个数M所对应的安全态势值作为输入值,根据该网络安全态势的预测模型对该输出层神经元个数N所对应的网络安全态势值进行预测;

其中所述步骤3包括步骤31、步骤32、步骤33、步骤34、步骤35、步骤36、步骤37、步骤38、步骤39:

步骤31,设定初始温度T_0,最小温度T_{min},温度T的迭代次数c,训练数据个数S,适应度阈值F和进化代数G,对所述训练数据进行长度为K的实数编码;

步骤32,计算编码后的训练数据的适应度值,适应度Fitness的计算公式为:

$$Fitness = \alpha(\sum_{i=1}^{n} abs(T_i - O_i)),$$

其中,α为函数系数,n为输出层神经元个数,T_i为输出层第i个神经元的安全态势值的期望输出值,O_i为输出层第i个神经元的安全态势值的预测输出值;

步骤33,使用适应度比例选择方法,选择出适应度不小于适应度阈值F的训练数据;

步骤34,根据交叉率交换训练数据的某些基因,将有益基因组合在一起;

步骤35,对训练数据的某些基因座上的基因值作变动,以维持该训练数据的多样性;

步骤36,对训练数据按照Metropolis准则进行接受;

步骤37,判断是否满足终止条件,若满足终止条件,则直接进入步骤39,若不满足终止条件,则进入下一步骤,其中,终止条件为到达所述训练数据预设定的最大进化次数,或连续多个新解未被接受,或达到预设最低温度T_{min};

步骤38,T_0向着T_{min}的方向逐渐降温,更新迭代次数,并转至步骤32,进行下一轮迭代;

步骤39,选择适应度最大的个体作为最具适应度的训练数据;

所述步骤36包括步骤361、步骤362:

步骤361,计算所述训练数据的进化代数G_t能量变化值$\Delta E=E(G_t)-E(G_{t-1})$,其中E(G)为能量的评价函数,取步骤32中所述的适应度为能量评价函数;

步骤362,若$\Delta E<0$则接受G_t作为新的训练数据,若$\Delta E>0$则以概率$p = e^{\frac{-\Delta E}{KT}}$接受$G_t$作为

新的训练数据；

所述步骤4包括步骤41、步骤42、步骤43、步骤44：

步骤41，利用所述步骤3的所述最具适应度的训练数据对应的权值和偏倚，对神经网络的权值和偏倚进行初始化赋值；

步骤42，该步骤包括判断所述反向传播神经网络训练次数是否满足迭代次数以及计算隐藏层输出、输出层输出；

步骤43，该步骤包括计算期望输出和输出层输出误差、判断输出层输出误差是否小于预设定阈值、计算隐藏层误差以及计算权值和偏倚的更新；

步骤44，根据步骤41-步骤43的计算与判断后，确立最终的网络安全态势的预测模型；

所述步骤42包括步骤421、步骤422、步骤423：

步骤421，判断所述神经网络的训练次数是否满足迭代次数，若满足迭代次数，则可确立预测模型，若不满足迭代次数，则进入下一步骤；

步骤422，隐藏层输出计算，输入的安全态势值向量$<ns_1,ns_2,\cdots,ns_M>$通过输入层不发生任何变化，即对于输入单元，它的输出O_j等于它的输入值ns_j，到达隐藏层后，隐藏层的净输入用其输入的线性组合计算$I_j=\sum_i w_{ij} O_i+\theta_j$，其中，j＝1,2,…h，h为隐藏层神经元个数，$w_{ij}$是由上一层的神经元i到神经元j连接的权重值，$O_i$是i的输出，$\theta_j$是j的偏倚，由神经元激励函数func计算得到隐藏层神经元j的输出O_j，$O_j=\mathrm{func}(I_j)$，$func(I_j)=\frac{1}{1+e^{-I_j}}$，j＝1,2,…h；

步骤423，输出层输出计算，根据隐藏层输出O_j，计算输出层输出O_k，

$O_k=\sum_j w_{jk} O_j+\theta_k$，

其中，k＝1,2,…n，n为输出层神经元个数，w_{jk}是由上一层的神经元j到神经元k的连接的权重值，O_j是j的输出，θ_k是k的偏倚。

2.如权利要求1所述的网络安全态势预测方法，其特征在于，所述步骤43具体为，

步骤431，输出层输出误差计算，对于输出层单元k误差Err_k的计算公式如下，

$Err_k=O_k(1-O_k)(T_k-O_k)$，

其中，T_k为期望输出的目标值，即真实获得的安全态势值ns_k，O_k为单元k输出的预测态势值ns'_k；

步骤432，判断输出层输出误差Err_k是否小于预设定阈值，若小于预设定阈值，则可确立预测模型，若不小于预设定阈值，则进入下一步骤；

步骤433，隐藏层误差计算，下一个较高层隐藏层单元j的误差Err_j的计算公式为如下，

$Err_j=O_j(1-O_j)\sum_k Err_k w_{jk}$；

步骤434，学习速率更新，假设可以利用上一轮t-1误差err_{t-1}和这一轮t误差err_t的增大减小的变化对lr_t进行微调，则lr_t的学习速率公式如下，

当$err_t>err_{t-1}$时，$lr_t = lr_{t-1} - lr_{t-1} \times \frac{t_{max}-t}{t_{max}} \times \left|\frac{err_t-err_{t-1}}{err_{t-1}}\right|$，

当$err_t<err_{t-1}$时，$lr_t = lr_{t-1} + lr_{t-1} \times \frac{t_{max}-t}{t_{max}} \times \left|\frac{err_t-err_{t-1}}{err_{t-1}}\right|$，

其中，t_{max}为预设定的最大迭代次数，t为当前进行的迭代轮数；

步骤435，权值更新，其更新公式如下，

$w_{ij}=w_{ij}+lr_t\ Err_j\ O_i$，$w_{jk}=w_{jk}+lr_t\ Err_k\ O_j$，

其中，lr_t为该轮迭代的学习速率；

步骤436，偏倚更新，输出层θ_k、隐藏层θ_j的更新公式如下，

$\theta_k=\theta_k+lr\ Err_k$，$\theta_j=\theta_j+lr\ Err_j$；

步骤437，进入步骤421，重复进行下一个周期。

3.一种网络安全态势预测系统，其特征在于，包括：

训练数据准备模块，用于将通过采集并融合局域网内资产、流量、入侵检测系统警报、漏洞数据而计算出的网络安全态势值序列集合作为训练数据；

初始化模块，用于对反向传播神经网络结构进行初始化，包括设定输入层神经元个数M、隐藏层神经元个数L和输出层神经元个数N；

模拟退火遗传算法优化模块，用于优化所述反向传播神经网络，包括对该训练数据进行长度为K的实数编码，其中，

K＝权值个数+偏倚个数＝(M*L+L*N)+(L+N)，

则每个编码后的该训练数据包含该反向传播神经网络的权值和偏倚信息，由适应度函数计算编码后的该训练数据的适应度值所确定的概率，找到最具适应度训练数据；

反向传播神经网络模型训练模块，用于训练网络安全态势的预测模型，包括将该最具适应度训练数据中的该输入层神经元个数M所对应的安全态势值作为输入值，该输出层神经元个数N所对应的安全态势值作为期望输出值，根据向前传递该输入值、向后传播该期望输出值来训练该反向传播神经网络，从而建立网络安全态势的预测模型；

模型预测模块，用于将该输入层神经元个数M所对应的安全态势值作为输入值，根据该网络安全态势的预测模型对该输出层神经元个数N所对应的网络安全态势值进行预测；

所述模拟退火遗传算法优化模块包括初始值设定模块、适应度值计算模块、选择操作模块、交叉操作模块、变异操作模块、Metropolis准则接受模块、终止条件判断模块、迭代次数更新模块、最具适应度的训练数据选择模块：

初始值设定模块，用于设定初始温度T_0、最小温度T_{min}、温度T的迭代次数c、训练数据个数S、适应度阈值F和进化代数G，并对所述训练数据进行长度为K的实数编码；

适应度值计算模块，用于计算编码后的训练数据的适应度值，适应度Fitness的计算公式为，

$$Fitness=\alpha(\sum_{i=1}^{n}abs(T_i-O_i)),$$

其中，α为函数系数，n为输出层神经元个数，T_i为输出层第i个神经元的安全态势值的期望输出值，O_i为输出层第i个神经元的安全态势值的预测输出值；

选择操作模块，用于选择操作使用适应度比例选择方法，选择出适应度不小于适应度阈值F的训练数据；

交叉操作模块，用于根据交叉率交换训练数据的某些基因，将有益基因组合在一起；

变异操作模块，用于对训练数据的某些基因座上的基因值作变动，以维持该训练数据的多样性；

Metropolis准则接受模块，用于对经过变异步骤后的训练数据按照Metropolis准则进

行接受；

终止条件判断模块，用于判断是否满足终止条件，若满足终止条件，则直接进入下述最具适应度的训练数据选择模块，若不满足终止条件，则进入下一步骤，其中，终止条件为到达所述训练数据预设定的最大进化次数，或连续多个新解未被接受，或达到预设最低温度T_{min}；

迭代次数更新模块，当T_0向着T_{min}的方向逐渐降温时，用于更新迭代次数，并转至所述适应度值计算模块，进行下一轮迭代；

最具适应度的训练数据选择模块，用于选择适应度最大的个体作为最具适应度的训练数据；

所述Metropolis准则接受模块包括能量变化值计算模块、接受模块：

能量变化值计算模块，用于计算所述训练数据的进化代数G_t能量变化值$\Delta E=E(G_t)-E(G_{t-1})$，其中E(G)为能量的评价函数，取所述适应度值计算模块中所述的适应度为能量评价函数；

接受模块，用于接受G_t作为新的训练数据，若$\Delta E<0$则接受G_t作为新的训练数据，若$\Delta E>0$则以概率$p=e^{\frac{-\Delta E}{KT}}$接受$G_t$作为新的训练数据；

所述反向传播神经网络模型训练模块包括权值和偏倚的初始化赋值模块、输入前向传递模块、输出后向传播模块、网络安全态势的预测模型确立模块：

权值和偏倚的初始化赋值模块，用于利用所述模拟退火遗传算法优化模块得到的所述最具适应度的训练数据对应的权值和偏倚，对神经网络的权值和偏倚进行初始化赋值；

输入前向传递模块，该模块包括用于判断所述反向传播神经网络训练次数是否满足迭代次数的训练次数判断模块以及隐藏层输出计算模块和输出层输出计算模块；

输出后向传播模块，该模块包括用于计算输出层输出误差的输出层误差计算模块、用于判断输出层输出误差是否小于预设定阈值的输出层误差判断模块、用于计算隐藏层误差的隐藏层误差计算模块、用于更新学习速率的学习速率更新模块以及权值更新计算模块和偏倚更新计算模块；

网络安全态势的预测模型确立模块，用于根据所述权值和偏倚的初始化赋值模块、所述输入前向传递模块和所述输出后向传播模块的计算与判断后，确立最终的网络安全态势的预测模型；

所述输入前向传递模块包括训练次数判断模块、隐藏层输出计算模块、输出层输出计算模块；

训练次数判断模块，用于判断所述神经网络的训练次数是否满足迭代次数，若满足迭代次数，则可确立预测模型，若不满足迭代次数，则进入下一步骤；

隐藏层输出计算模块，用于计算隐藏层输出，具体为，输入的安全态势值向量$<ns_1,ns_2,\cdots,ns_M>$通过输入层不发生任何变化，即对于输入单元，它的输出O_j等于它的输入值ns_j，到达隐藏层后，隐藏层的净输入用其输入的线性组合计算$I_j=\sum_i w_{ij} O_i+\theta_j$，其中，j=1,2,…h，h为隐藏层神经元个数，$w_{ij}$是由上一层的神经元i到神经元j连接的权重值，$O_i$是i的输出，$\theta_j$是j的偏倚，由神经元激励函数func计算得到隐藏层神经元j的输出O_j，$O_j=func(I_j)$，$func(I_j)=\frac{1}{1+e^{-I_j}}$，j=1,2,…h；

输出层输出计算模块，用于计算输出层输出，具体为，根据隐藏层输出O_j，计算输出层输出O_k，

$O_k=\Sigma_j\ w_{jk}\ O_j+\theta_k$，

其中，k＝1,2,…n，n为输出层神经元个数，w_{jk}是由上一层的神经元j到神经元k的连接的权重值，O_j是j的输出，θ_k是k的偏倚。

4.如权利要求3所述的网络安全态势预测系统，其特征在于，所述输出后向传播模块进一步包括，

输出层误差计算模块，用于计算输出层输出误差，具体为，对于输出层单元k误差Err_k的计算公式如下，

$Err_k=O_k(1-O_k)(T_k-O_k)$，

其中，T_k为期望输出的目标值，即真实获得的安全态势值ns_k，O_k为单元k输出的预测态势值ns'_k；

输出层误差判断模块，用于判断输出层输出误差Err_k是否小于预设定阈值，若小于预设定阈值，则可确立预测模型，若不小于预设定阈值，则进入下一步骤；

隐藏层误差计算模块，用于计算隐藏层误差，下一个较高层隐藏层单元j的误差Err_j的计算公式为如下，

$Err_j=O_j(1-O_j)\Sigma_k\ Err_k\ w_{jk}$；

学习速率更新模块，用于更新学习速率，假设可以利用上一轮t-1误差err_{t-1}和这一轮t误差err_t的增大减小的变化对lr_t进行微调，则lr_t的学习速率公式如下，

当$err_t>err_{t-1}$时，$$lr_t = lr_{t-1} - lr_{t-1} \times \frac{t_{max}-t}{t_{max}} \times \left|\frac{err_t-err_{t-1}}{err_{t-1}}\right|,$$

当$err_t<err_{t-1}$时，$$lr_t = lr_{t-1} + lr_{t-1} \times \frac{t_{max}-t}{t_{max}} \times \left|\frac{err_t-err_{t-1}}{err_{t-1}}\right|,$$

其中，t_{max}为预设定的最大迭代次数，t为当前进行的迭代轮数；

权值更新计算模块，用于计算权值更新，其更新公式如下，

$w_{ij}=w_{ij}+lr_t\ Err_j\ O_i$，$w_{jk}=w_{jk}+lr_t\ Err_k\ O_j$，

其中，lr_t为该轮迭代的学习速率；

偏倚更新计算模块，用于计算偏倚更新，输出层θ_k、隐藏层θ_j的更新公式如下，

$\theta_k=\theta_k+lr\ Err_k$，$\theta_j=\theta_j+lr\ Err_j$；

最后，进入所述训练次数判断模块，重复进行下一个周期。

一种网络安全态势预测方法及系统

技术领域

[0001] 本发明涉及网络安全技术领域，特别是涉及一种网络安全态势预测方法及系统。

背景技术

[0002] 计算机网络是通信技术和计算机技术发展到一定程度后相结合的产物，在高度发展的网络技术为人们带来快速便捷的信息交互的同时，网络的恶意攻击和窃取等行为也愈演愈烈。攻击者利用网络的快速传播性和广泛互联性，大肆地破坏网络的基本性能、侵害用户的合法权益，威胁社会和国家的安全利益，因而使传统意义上的网络安全措施面临严峻的考验。随着网络入侵行为向着多元化、规模化、复杂化、持续化等趋势发展，安全管理者越来越希望更好地了解其监管的网络当前时刻和未来时刻的安全健康状态，以便及时发现问题、采取预警措施，因此，网络安全态势感知技术应运而生。近年来，网络安全态势感知成为当前网络安全界研究的热点，这项研究取得的成果，在提高网络的监控、应急响应能力和预测网络的安全发展趋势等方面都将起到重大的推动作用。

[0003] 网络安全态势预测是指依据预测模型对网络安全发展趋势，即恶意攻击和窃取行为对未来网络安全所造成的影响进行判定。为制定应急处理方案、提高网络响应能力作有力基础保障，如何构建有效的网络安全态势的预测模型具有很大的难度，在以往的预测方法中，一般采用先验知识构建态势预测模型，也可以采用数据挖掘的方法从大量网络安全态势的历史数据中学习模型来预测网络的发展趋势，例如支持向量机、时间序列法等，但是由于网络安全态势的发展具有非线性、随机性和不确定性，以上这些方法应用时都具有一定的局限性。

[0004] 人工神经网络是对自然的神经网络的模仿，可以普适性的解决复杂 的包含大量相互相关的变量预测、回归和分类等问题。神经网络具有对噪声数据的高承受力，神经网络学习对于训练数据中的错误表现有较高的健壮性，在缺乏属性与类之间的联系的知识时可以使用神经网络，尤其适合连续值的输入和输出。在训练神经网络时，可能花费一些时间，但一旦训练完毕，进行新数据预测计算时很快，因此，神经网络的训练已被成功应用到很多领域。

[0005] 反向传播神经网络是一种由反向传播算法训练形成的具有输入向前传递，误差向后传播特点的多层前馈神经网络。在向前传递中，输入信号由输入层经隐藏层逐层处理计算，直到输出层输出；输出值与期望值的误差后向传播，将误差信号反向由输出层通过隐藏层处理后向输入层传播，根据预测误差调整网络的权值和偏倚，误差依据梯度下降算法分摊给各层的所有单元，从而获得各单元的误差信号来修正各神经元权值和偏倚。继续重复输入向前传递、误差向后传播的过程，使得反向传播神经网络预测输出不断逼近期望输出。通过以达到预定的可接受误差或是达到设定的学习次数为终止条件来终止学习训练过程。

[0006] 目前，已有如下网络安全态势的预测方法：

[0007] 如发明名称为“一种网络安全态势预测方法”，公开了网络信息安全技术领域中的一种网络安全态势预测方法。该方法使用灰色聚类分析方法分析每种网络安全威胁的危害

程度，进而构造出层次化的网络安全态势指标体系，得到每个时间监测点的网络安全态势值并构造成时间序列，将其构造成训练样本集，利用集成学习Boosting算法对训练样本集进行迭代训练得到满足误差要求的弱学习机序列；再利用对弱学习机序列加权求和的方法得到强学习机；利用强学习机完成未来时间监测点的网络安全态势值预测。该发明在降低网络安全态势值预测误差方面，有较好的适应性和较低的预测误差。而该发明使用的方法是Boosting，该方法过于依赖数据和弱学习机，对数据噪声很敏感，而且如果弱学习机过弱也不能够达到较高的预测精度。

[0008] 如发明名称为“网络安全态势预测的高斯过程回归方法”，该发明公开了网络信息安全技术领域的网络安全态势预测的高斯过程回归方法。该发明使用层次分析法构造出层次化网络安全态势评价指标体系，以该体系分析各种网络安全威胁对网络安全态势的危害程度，进而计算出各个时间监测点的网络安全态势值并构造成时间序列，将其构造成训练样本集，利用高斯过程回归对训练样本集进行迭代训练得到满足误差要求的预测模型，在训练过程中利用粒子群算法动态搜索高斯过程回归的最优训练参数以降低预测误差，最后利用预测模型完成未来时间监测点的网络安全态势值预测。该发明能在降低网络安全态势预测误差方面，有较好的适应性和较低的预测误差。该专利使用的方法是粒子群优化的高斯过程回归，该方法计算量大，需要效率较高的协方差求逆计算方法或训练集选择方法；此外，该方法的原理中假设噪声必须满足高斯分布，所以实际运用中还需要对预测空间的数据进行相对复杂的变换以满足该假设。

发明内容

[0009] 为了解决上述问题，本发明的目的在于提供一种网络安全态势预测方法及系统，该方法采用模拟退火遗传算法优化反向传播神经网络的方法来训练网络安全态势的预测模型，并通过该预测模型对未来时刻的网络安全态势进行预测，从而克服已有网络安全态势预测方法的缺陷，提高网络安全态势预测方法的收敛速度，降低训练时间和预测误差。

[0010] 为实现上述目的，本发明所提出的网络安全态势预测方法，其特征在于，包括以下步骤：

[0011] 步骤1，将通过采集并融合局域网内资产、流量、入侵检测系统警报、漏洞数据而计算出的多个网络安全态势值作为训练数据；

[0012] 步骤2，对反向传播神经网络结构进行初始化，包括设定输入层神经元个数M、隐藏层神经元个数L和输出层神经元个数N；

[0013] 步骤3，对该训练数据进行长度为K的实数编码，其中，

[0014] K=权值个数+偏倚个数=(M*L+L*N)+(L+N)，

[0015] 则每个编码后的该训练数据包含该反向传播神经网络的权值和偏倚信息，由适应度函数计算编码后的该训练数据的适应度值所确定的概率，找到最具适应度训练数据；

[0016] 步骤4，将该最具适应度训练数据中的该输入层神经元个数M所对应的安全态势值作为输入值，该输出层神经元个数N所对应的安全态势值作为期望输出值，根据向前传递该输入值、向后传播该期望输出值来训练该反向传播神经网络，从而建立网络安全态势的预测模型；

[0017] 步骤5，将该输入层神经元个数M所对应的安全态势值作为输入值，根据该网络安

全态势的预测模型对该输出层神经元个数N所对应的网络安全态势值进行预测。

[0018] 本发明所提出的网络安全态势预测方法，其特征在于，所述步骤3进一步包括以下步骤：

[0019] 步骤31，设定初始温度T_0，最小温度T_{min}，温度T的迭代次数c，训练数据个数S，适应度阈值F和进化代数G，对所述训练数据进行长度为K的实数编码；

[0020] 步骤32，计算编码后的训练数据的适应度值，适应度Fitness的计算公式为：

[0021] $$Fitness = \alpha(\sum_{i=1}^{n} abs(T_i - O_i)),$$

[0022] 其中，α为函数系数，n为输出层神经元个数，T_i为输出层第i个神经元的安全态势值的期望输出值，O_i为输出层第i个神经元的安全态势值的预测输出值；

[0023] 步骤33，使用适应度比例选择方法，选择出适应度不小于适应度阈值F的训练数据；

[0024] 步骤34，根据交叉率交换训练数据的某些基因，将有益基因组合在一起；

[0025] 步骤35，对训练数据的某些基因座上的基因值作变动，以维持该训练数据的多样性；

[0026] 步骤36，对训练数据按照Metropolis准则进行接受；

[0027] 步骤37，判断是否满足终止条件，若满足终止条件，则直接进入步骤39，若不满足终止条件，则进入下一步骤，其中，终止条件为到达所述训练数据预设定的最大进化次数，或连续多个新解未被接受，或达到预设最低温度T_{min}；

[0028] 步骤38，T_0向着T_{min}的方向逐渐降温，更新迭代次数，并转至步骤32，进行下一轮迭代；

[0029] 步骤39，选择适应度最大的个体作为最具适应度的训练数据。

[0030] 本发明所提出的网络安全态势预测方法，其特征在于，步骤36进一步包括以下步骤：

[0031] 步骤361，计算所述训练数据的进化代数G_t能量变化值$\Delta E=E(G_t)-E(G_{t-1})$，其中E(G)为能量的评价函数，取步骤32中所述的适应度为能量评价函数；

[0032] 步骤362，若$\Delta E<0$则接受G_t作为新的训练数据，若$\Delta E>0$则以概率$p = e^{\frac{-\Delta E}{KT}}$接受$G_t$作为新的训练数据。

[0033] 本发明所提出的网络安全态势预测方法，其特征在于，步骤4进一步包括以下步骤：

[0034] 步骤41，利用所述步骤3的所述最具适应度的训练数据对应的权值和偏倚，对神经网络的权值和偏倚进行初始化赋值；

[0035] 步骤42，该步骤包括判断所述反向传播神经网络训练次数是否满足迭代次数以及计算隐藏层输出、输出层输出；

[0036] 步骤43，该步骤包括计算期望输出和输出层输出误差、判断输出层输出误差是否小于预设定阈值、计算隐藏层误差以及计算权值和偏倚的更新；

[0037] 步骤44，根据步骤41-步骤43的计算与判断后，确立最终的网络安全态势的预测模

型。

[0038] 本发明所提出的网络安全态势预测方法，其特征在于，所述步骤42具体为，

[0039] 步骤421，判断所述神经网络的训练次数是否满足迭代次数，若满足迭代次数，则可确立预测模型，若不满足迭代次数，则进入下一步骤；

[0040] 步骤422，隐藏层输出计算，输入的安全态势值向量＜ns_1,ns_2,…,ns_M＞通过输入层不发生任何变化，即对于输入单元，它的输出O_j等于它的输入值ns_j，到达隐藏层后，隐藏层的净输入用其输入的线性组合计算$I_j=\Sigma_i w_{ij}O_i+\theta_j$，其中，j=1,2,…h，h为隐藏层神经元个数，$w_{ij}$是由上一层的神经元i到神经元j连接的权重值，$O_i$是 i的输出，$\theta_j$是j的偏倚，由神经元激励函数func计算得到隐藏层神经元j的输出O_j，$O_j=func(I_j)$，j=1,2,…h；

[0041] 步骤423，输出层输出计算，根据隐藏层输出O_j，计算输出层输出O_k，

[0042] $O_k=\Sigma_j w_{jk}O_j+\theta_k$，

[0043] 其中，k=1,2,…n，n为输出层神经元个数，w_{jk}是由上一层的神经元j到神经元k的连接的权重值，O_j是j的输出，θ_k是k的偏倚；

[0044] 本发明所提出的网络安全态势预测方法，其特征在于，所述步骤43具体为，

[0045] 步骤431，输出层输出误差计算，对于输出层单元k误差Err_k的计算公式如下，

[0046] $Err_k=O_k(1-O_k)(T_k-O_k)$，

[0047] 其中，T_k为期望输出的目标值，即真实获得的安全态势值ns_k，O_k为单元k输出的预测态势值ns_k'；

[0048] 步骤432，判断输出层输出误差Err_k是否小于预设定阈值，若小于预设定阈值，则可确立预测模型，若不小于预设定阈值，则进入下一步骤；

[0049] 步骤433，隐藏层误差计算，下一个较高层隐藏层单元j的误差Err_j的计算公式为如下，

[0050] $Err_j=O_j(1-O_j)\Sigma_k Err_k w_{jk}$；

[0051] 步骤434，学习速率更新，假设可以利用上一轮t-1误差err_{t-1}和这一轮t误差err_t的增大减小的变化对lr_t进行微调，则lr_t的学习速率公式如下，

[0052] 当err_t＞err_{t-1}时，$$lr_t = lr_{t-1} - lr_{t-1} \times \frac{t_{max}-t}{t_{max}} \times \left|\frac{err_t - err_{t-1}}{err_{t-1}}\right|,$$

[0053] 当err_t＜err_{t-1}时，$$lr_t = lr_{t-1} + lr_{t-1} \times \frac{t_{max}-t}{t_{max}} \times \left|\frac{err_t - err_{t-1}}{err_{t-1}}\right|,$$

[0054] 其中，t_{max}为预设定的最大迭代次数，t为当前进行的迭代轮数；

[0055] 步骤435，权值更新，其更新公式如下，

[0056] $w_{ij}=w_{ij}+lr_t Err_j O_i$，$w_{jk}=w_{jk}+lr_t Err_k O_j$，

[0057] 其中，lr_t为该轮迭代的学习速率；

[0058] 步骤436，偏倚更新，输出层θ_k、隐藏层θ_j的更新公式如下，

[0059] $\theta_k=\theta_k+lrErr_k$，$\theta_j=\theta_j+lrErr_j$；

[0060] 步骤437，进入步骤421，重复进行下一个周期。

[0061] 本发明还涉及一种网络安全态势预测系统，其特征在于，包括：

[0062] 训练数据准备模块，用于将通过采集并融合局域网内资产、流量、入侵检测系统警报、漏洞数据而计算出的网络安全态势值序列集合作为训练数据；

[0063] 初始化模块，用于对反向传播神经网络结构进行初始化，包括设定输入层神经元个数M、隐藏层神经元个数L和输出层神经元个数N；

[0064] 模拟退火遗传算法优化模块，用于优化所述反向传播神经网络，包括对该训练数据进行长度为K的实数编码，其中，

[0065] K=权值个数+偏倚个数=(M*L+L*N)+(L+N)，

[0066] 则每个编码后的该训练数据包含该反向传播神经网络的权值和偏倚信息，由适应度函数计算编码后的该训练数据的适应度值所确定的概率，找到最具适应度训练数据；

[0067] 反向传播神经网络模型训练模块，用于训练网络安全态势的预测模型，包括将该最具适应度训练数据中的该输入层神经元个数M所对应的安全态势值作为输入值，该输出层神经元个数N所对应的安全态势值作为期望输出值，根据向前传递该输入值、向后传播该期望输出值来训练该反向传播神经网络，从而建立网络安全态势的预测模型；

[0068] 模型预测模块，用于将该输入层神经元个数M所对应的安全态势值作为输入值，根据该网络安全态势的预测模型对该输出层神经元个数N所对应的网络安全态势值进行预测。

[0069] 本发明所提出的网络安全态势预测系统，其特征在于，所述模拟退火遗传算法优化模块进一步包括以下模块：

[0070] 初始值设定模块，用于设定初始温度T_0、最小温度T_{min}、温度T的迭代次数c、训练数据个数S、适应度阈值F和进化代数G，并对所述训练数据进行长度为K的实数编码；

[0071] 适应度值计算模块，用于计算编码后的训练数据的适应度值，适应度Fitness的计算公式为，

[0072] $$Fitness = \alpha(\sum_{i=1}^{n} abs(T_i - O_i)),$$

[0073] 其中，α为函数系数，n为输出层神经元个数，T_i为输出层第i个神经元的安全态势值的期望输出值，O_i为输出层第i个神经元的安全态势值的预测输出值；

[0074] 选择操作模块，用于选择操作使用适应度比例选择方法，选择出适应度不小于适应度阈值F的训练数据；

[0075] 交叉操作模块，用于根据交叉率交换训练数据的某些基因，将有益基因组合在一起；

[0076] 变异操作模块，用于对训练数据的某些基因座上的基因值作变动，以维持该训练数据的多样性；

[0077] Metropolis准则接受模块，用于对经过变异步骤后的训练数据按照Metropolis准则进行接受；

[0078] 终止条件判断模块，用于判断是否满足终止条件，若满足终止条件，则直接进入下述最具适应度的训练数据选择模块，若不满足终止条件，则进入下一步骤，其中，终止条件为到达所述训练数据预设定的最大进化次数，或连续多个新解未被接受，或达到预设最低温度T_{min}；

[0079] 迭代次数更新模块，当T_0向着T_{min}的方向逐渐降温时，用于更新迭代次数，并转至所述适应度值计算模块，进行下一轮迭代；

[0080] 最具适应度的训练数据选择模块，用于选择适应度最大的个体作为最具适应度的训练数据。

[0081] 本发明所提出的网络安全态势预测系统，其特征在于，所述Metropolis准则接受模块进一步包括以下模块：

[0082] 能量变化值计算模块，用于计算所述训练数据的进化代数G_t能量变化值$\Delta E=E(G_t)-E(G_{t-1})$，其中E(G)为能量的评价函数，取所述适应度值计算模块中所述的适应度为能量评价函数；

[0083] 接受模块，用于接受G_t作为新的训练数据，若$\Delta E<0$则接受G_t作为新的训练数据，若$\Delta E>0$则以概率$p=e^{\frac{-\Delta E}{KT}}$接受$G_t$作为新的训练数据。

[0084] 本发明所提出的网络安全态势预测系统，其特征在于，所述反向传播神经网络模型训练模块进一步包括以下模块：

[0085] 权值和偏倚的初始化赋值模块，用于利用所述模拟退火遗传算法优化模块得到的所述最具适应度的训练数据对应的权值和偏倚，对神经网络的权值和偏倚进行初始化赋值；

[0086] 输入前向传递模块，该模块包括用于判断所述反向传播神经网络训练次数是否满足迭代次数的训练次数判断模块以及隐藏层输出计算模块和输出层输出计算模块；

[0087] 输出后向传播模块，该模块包括用于计算输出层输出误差的输出层误差计算模块、用于判断输出层输出误差是否小于预设定阈值的输出层误差判断模块、用于计算隐藏层误差的隐藏层误差计算模块、用于更新学习速率的学习速率更新模块以及权值更新计算模块和偏倚更新计算模块；

[0088] 网络安全态势的预测模型确立模块，用于根据所述权值和偏倚的初始化赋值模块、所述输入前向传递模块和所述输出后向传播模块的计算与判断后，确立最终的网络安全态势的预测模型。

[0089] 本发明所提出的网络安全态势预测系统，其特征在于，所述输入前向传递模块进一步包括，

[0090] 训练次数判断模块，用于判断所述神经网络的训练次数是否满足迭代次数，若满足迭代次数，则可确立预测模型，若不满足迭代次数，则进入下一步骤；

[0091] 隐藏层输出计算模块，用于计算隐藏层输出，具体为，输入的安全态势值向量$<ns_1,ns_2,\cdots,ns_M>$通过输入层不发生任何变化，即对于输入单元，它的输出O_j等于它的输入值ns_j，到达隐藏层后，隐藏层的净输入用其输入的线性组合计算$I_j=\Sigma_i w_{ij}O_i+\theta_j$，其中，j=1,2,…h，h为隐藏层神经元个数，$w_{ij}$是由上一层的神经元i到神经元j连接的权重值，$O_i$是i的输出，$\theta_j$是j的偏倚，由神经元激励函数func计算得到隐藏层神经元j的输出O_j，$O_j=func(I_j)$，j=1,2,…h；

[0092] 输出层输出计算模块，用于计算输出层输出，具体为，根据隐藏层输出O_j，计算输出层输出O_k，

[0093] $O_k=\Sigma_j w_{jk}O_j+\theta_k$，

[0094] 其中，k=1,2,…n，n为输出层神经元个数，w_{jk}是由上一层的神经 元j到神经元k的连接的权重值，O_j是j的输出，θ_k是k的偏倚；

[0095] 本发明所提出的网络安全态势预测系统，其特征在于，所述输出后向传播模块进一步包括，

[0096] 输出层误差计算模块，用于计算输出层输出误差，具体为，对于输出层单元k误差Err_k的计算公式如下，

[0097] $Err_k=O_k(1-O_k)(T_k-O_k)$，

[0098] 其中，T_k为期望输出的目标值，即真实获得的安全态势值ns_k，O_k为单元k输出的预测态势值ns_k'；

[0099] 输出层误差判断模块，用于判断输出层输出误差Err_k是否小于预设定阈值，若小于预设定阈值，则可确立预测模型，若不小于预设定阈值，则进入下一步骤；

[0100] 隐藏层误差计算模块，用于计算隐藏层误差，下一个较高层隐藏层单元j的误差Err_j的计算公式为如下，

[0101] $Err_j=O_j(1-O_j)\sum_k Err_k w_{jk}$；

[0102] 学习速率更新模块，用于更新学习速率，假设可以利用上一轮t-1误差err_{t-1}和这一轮t误差err_t的增大减小的变化对lr_t进行微调，则lr_t的学习速率公式如下，

[0103] 当$err_t > err_{t-1}$时，$lr_t = lr_{t-1} - lr_{t-1} \times \frac{t_{max}-t}{t_{max}} \times \left|\frac{err_t - err_{t-1}}{err_{t-1}}\right|$，

[0104] 当$err_t < err_{t-1}$时，$lr_t = lr_{t-1} + lr_{t-1} \times \frac{t_{max}-t}{t_{max}} \times \left|\frac{err_t - err_{t-1}}{err_{t-1}}\right|$，

[0105] 其中，t_{max}为预设定的最大迭代次数，t为当前进行的迭代轮数；

[0106] 权值更新计算模块，用于计算权值更新，其更新公式如下，

[0107] $w_{ij}=w_{ij}+lr_t Err_j O_i$，$w_{jk}=w_{jk}+lr_t Err_k O_j$，

[0108] 其中，lr_t为该轮迭代的学习速率；

[0109] 偏倚更新计算模块，用于计算偏倚更新，输出层θ_k、隐藏层θ_j的更新公式如下，

[0110] $\theta_k=\theta_k+lrErr_k$，$\theta_j=\theta_j+lrErr_j$；

[0111] 最后，进入所述训练次数判断模块，重复进行下一个周期。

[0112] 本发明提供了网络安全技术领域中的一种网络安全态势预测方法，该方法通过将一定时间间隔获得的网络安全态势值序列集合作为预测 模型的训练数据，应用模拟退火遗传算法优化反向传播神经网络来训练出网络安全态势的预测模型，最后利用预测模型对未来时刻的网络安全态势值进行预测。本发明的应用效果表明，该方法的收敛速度得到提高、降低了训练时间和预测误差。

[0113] 相较于其他网络安全态势预测方法，本发明提出的技术方案，具有以下优势：第一，具有对噪声数据的高承受力，对于训练数据中的错误表现有较高的健壮性；第二，在缺乏属性与类之间的联系的知识时可以使用，尤其适合连续值的输入和输出，使用者无需知道未来时刻态势值与历史态势值间的函数关系，即可以在一定的误差允许范围，来逼近这一函数映射；第三，在训练预测模型时，可能花费一些时间，但一旦训练完毕，进行新数据预测计算时是很快的；第四，通过应用误差和迭代次数来调节学习率、采用模拟退火遗传算法优化等措施，在训练时间未增加的条件下，可以提高预测方法的收敛速度、降低误差。

附图说明

[0114] 图1为本发明的网络安全态势预测方法的流程图；

[0115] 图2是本发明的网络安全态势预测方法的简要流程说明图；

[0116] 图3是本发明的网络安全态势预测方法的具体流程说明图；

[0117] 图4是本发明的网络安全态势预测仿真图。

具体实施方式

[0118] 为了使本发明的目的、技术方案及优点更加清楚明白，以下对本发明提出的网络安全态势预测方法进行进一步详细说明。

[0119] 本发明所提出的网络安全态势预测方法，如图2所示，使用模拟退火遗传算法优化反向传播神经网络，然后使用历史时刻的网络安全态势值训练反向传播神经网络，最后使用训练好的反向传播神经网络进行下一时刻网络安全态势值的预测，该方法的操作步骤如图1所示，具体为：

[0120] 步骤1，使用发明人开发的网络安全态势感知系统，将通过采集并融合局域网内资产、流量、入侵检测系统警报、漏洞数据而计算出的历史安全态势值序列集合作为预测模型的训练数据；

[0121] 步骤2，对反向传播神经网络结构进行初始化，包括设定输入层神经元个数M、隐藏层神经元个数L和输出层神经元个数N；

[0122] 步骤3，对该训练数据进行编码长度为K的实数编码，其中，

[0123] K=权值个数+偏倚个数=(M*L+L*N)+(L+N)，则每个编码后的该训练数据包含该反向传播神经网络的权值和偏倚信息，由适应度函数计算编码后的该训练数据的适应度值所确定的概率大小来进行选择操作、交叉操作和变异操作，找到最具适应度的训练数据；

[0124] 步骤4，将该最具适应度的训练数据中的该输入层神经元个数M所对应的安全态势值作为输入值，该输出层神经元个数N所对应的安全态势值作为期望输出值，根据向前传递该输入值、向后传播该期望输出值来训练所述反向传播神经网络，从而建立网络安全态势的预测模型，其中，M、N为大于1的自然数，当满足训练终止条件时，预测模型的学习训练完成；

[0125] 步骤5，将该输入层神经元个数M所对应的安全态势值作为输入值，根据训练出的预测模型对未来该输出层神经元个数N所对应的网络安全态势值进行预测。

[0126] 根据本发明，其中，步骤3进一步包括以下步骤：

[0127] 步骤31，设定初始温度T_0，最小温度T_{min}，温度T的迭代次数c，训练数据个数S、适应度阈值F和进化代数G，对所述训练数据进行长度为K的实数编码，编码长度K=权值个数+偏倚个数=(M*L+L*N)+(L+N)，编码后的训练数据由输入层与隐藏层神经元之间、隐藏层与输出层神经元之间的连接权值w_{ij}、w_{jk}，隐藏层神经元、输出层神经元的偏倚θ_k、θ_n，共计四部分组成，因此个体包含了神经网络的全部权值和偏倚，其中，T_{min}，c，S，G、F一般根据经验值来设定；

[0128] 步骤32，计算编码后的训练数据的适应度值，适应度Fitness的计算公式为：

[0129] $$Fitness = \alpha(\sum_{i=1}^{n} abs(T_i - O_i))$$

[0130] 其中，α为函数系数，n为输出层神经元个数，T_i为输出层第i个神经元的安全态势值的期望输出值，O_i为输出层第i个神经元的安全态势 值的预测输出值；

[0131] 步骤33，选择操作，选择操作使用的是适应度比例选择方法，即轮盘赌选择进行，选择适应度不小于适应度阈值F的训练数据，在该方法中，每个编码后的训练数据i的选择概率和其适应度值成比例，该选择概率反映了数据i的适应度在整个群体的个体适应度总和中所占的比例，个体适应度越大，其被选择的概率就越高、反之亦然，数据i的选择概率p_i的公式为，

[0132] $$p_i = \frac{Fitness_i}{\sum_{j=1}^{S} Fitness_j}$$

[0133] 其中，$Fitness_i$为数据i的适应度值，S为训练数据个数；

[0134] 步骤34，交叉操作，依据实数交叉法进行，根据交叉率将种群中的两个数据随机地交换某些基因，能够产生新的基因组合，期望将有益基因组合在一起，第s个数据a_s和第l个数据在第a_l位置交叉如下：

[0135] $$\begin{cases} a_{sj} = a_{sj}(1-b) + a_{lj}b \\ a_{lj} = a_{lj}(1-b) + a_{sj}b \end{cases}$$

[0136] 其中，b为[0,1]之间的随机数；

[0137] 步骤35，变异操作，依据实数变异法进行，对群体中的数据串的某些基因座上的基因值作变动，以使遗传算法具有局部的随机搜索能力，并可维持群体多样性，第i个数据的第j位置a_{ij}进行变异如下：

[0138] $$a_{ij} = \begin{cases} a_{ij} + (a_{ij} - a_{max}) * f(G) & \gamma_1 \geq 0.5 \\ a_{ij} + (a_{min} - a_{ij}) * f(G) & \gamma_1 < 0.5 \end{cases}$$

[0139] 其中，γ_1为[0,1]间的随机数，a_{min}为权值或偏倚a_{ij}的下界，a_{max}为权值或偏倚a_{ij}的上界，$f(G) = \gamma_2(1 - \frac{G}{G_{max}})$，$\gamma_2$=0.1，G为当前进化代数，$G_{max}$是预设定的最大进化代数，并对经过变异步骤后的训练数据按照模拟退火算法准则进行接受；

[0140] 步骤36，对经过变异步骤后的训练数据按照Metropolis准则进行接受，如存在不被接受的数据，则直接舍弃该数据，具体为：

[0141] 步骤361，计算训练数据的进化代数G_t能量变化值ΔE=E(G_t)-E(G_{t-1})，其中E(G)为能量的评价函数，取步骤32中所述的适应度为能量评价函数，即$E(G_t) = \frac{1}{d}\sum_{i=1}^{d} Fitness_{x_i}$，其中$x_i$为$G_t$中的编码后的训练数据，$Fitness_{x_i}$为$x_i$的适应度，d为$G_t$的训练数据个数；

[0142] 步骤362，若ΔE＜0则接受G_t作为新的训练数据，若ΔE＞0则以概率$p = e^{\frac{-\Delta E}{KT}}$接受$G_t$作为新的训练数据；

[0143] 步骤37，判断是否满足终止条件，若满足终止条件，则直接进入步骤39，若不满足终止条件，则进入下一步骤，其中，终止条件为到达所述训练数据预设定的最大进化次数，或连续多个新解未被接受，或达到预设最低温度T_{min}；

[0144] 步骤38，初始温度T_0向着T_{min}的方向逐渐降温，更新迭代次数，并转至步骤32，进行下一轮迭代；

[0145] 步骤39，选择最终得到的子代中适应度最大的个体作为最具适应度的训练数据。

[0146] 根据本发明，其中，步骤4进一步包括以下步骤：

[0147] 步骤41，利用步骤3得到的最具适应度的训练数据对应的权值和偏倚，对神经网络的权值和偏倚进行初始化赋值；

[0148] 步骤42，该步骤包括判断神经网络训练次数是否满足迭代次数以及隐藏层输出、输出层输出的计算。

[0149] 步骤421，判断所述神经网络的训练次数是否满足迭代次数，若满足迭代次数，则可确立预测模型，若不满足迭代次数，则进入下一步骤；

[0150] 步骤422，隐藏层输出计算，输入的安全态势值向量＜ns_1，ns_2，…，ns_M＞通过输入层不发生任何变化，即对于输入单元，它的输出O_j等于它的输入值ns_j，到达隐藏层后，隐藏层的净输入用其输入的线性组合计算$I_j=\Sigma_i w_{ij}O_i+\theta_j$，其中，j=1，2，…h，h为隐藏层神经元个数，$w_{ij}$是由上一层的神经元i到神经元j连接的权重值，$O_i$是i的输出，$\theta_j$是j的偏倚，由神经元激励函数func计算得到隐藏层神经元j的输出O_j，$O_j=func(I_j)$，j=1，2，…h；

[0151] 步骤423，输出层输出计算，根据隐藏层输出O_j，计算输出层输出O_k，

[0152] $O_k=\Sigma_j w_{jk}O_j+\theta_k$，

[0153] 其中，k=1，2，…n，n为输出层神经元个数，w_{jk}是由上一层的神经 元j到神经元k的连接的权重值，O_j是j的输出，θ_k是k的偏倚；

[0154] 步骤43，该步骤包括期望输出和输出层输出误差的计算、输出层输出误差的判断、隐藏层误差的计算以及权值和偏倚的更新计算，具体为：

[0155] 步骤431，输出层输出误差计算，对于输出层单元k误差Err_k的计算公式如下，

[0156] $Err_k=O_k(1-O_k)(T_k-O_k)$，

[0157] 其中，T_k为期望输出的目标值，即真实获得的安全态势值ns_k，O_k为单元k输出的预测态势值ns_k'；

[0158] 步骤432，判断输出层输出误差Err_k是否小于预设定阈值，若小于预设定阈值，则可确立预测模型，若不小于预设定阈值，则进入下一步骤；

[0159] 步骤433，隐藏层误差计算，下一个较高层隐藏层单元j的误差Err_j的计算公式为如下，

[0160] $Err_j=O_j(1-O_j)\Sigma_k Err_k w_{jk}$；

[0161] 步骤434，学习速率更新，假设可以利用上一轮t-1误差err_{t-1}和这一轮t误差err_t的增大减小的变化对lr_t进行微调，则lr_t的学习速率公式如下，

[0162] 当err_t＞err_{t-1}时，$$lr_t = lr_{t-1} - lr_{t-1} \times \frac{t_{max}-t}{t_{max}} \times \left|\frac{err_t - err_{t-1}}{err_{t-1}}\right|$$

[0163] 当err_t＜err_{t-1}时，$$lr_t = lr_{t-1} + lr_{t-1} \times \frac{t_{max}-t}{t_{max}} \times \left|\frac{err_t - err_{t-1}}{err_{t-1}}\right|$$

[0164] 其中，t_{max}为预设定的最大迭代次数，t为当前进行的迭代轮数；

[0165] 步骤435，权值更新，其更新公式如下，

[0166] $w_{ij}=w_{ij}+lr_t Err_j O_i$，$w_{jk}=w_{jk}+lr_t Err_k O_j$

[0167] 其中，lr_t为该轮迭代的学习速率；

[0168] 步骤436，偏倚更新，输出层θ_k、隐藏层θ_j的更新公式如下，

[0169] $\theta_k=\theta_k+lrErr_k$，$\theta_j=\theta_j+lrErr_j$；

[0170] 步骤437，进入步骤421，重复进行下一个周期。

[0171] 经过步骤41-步骤43的一系列计算与判断后，确立最终的网络安全态势的预测模型。

[0172] 本发明提出的网络安全态势预测系统，包括：

[0173] 训练数据准备模块，用于将通过采集并融合局域网内资产、流量、入侵检测系统警报、漏洞数据而计算出的网络安全态势值序列集合作为训练数据；

[0174] 初始化模块，用于对反向传播神经网络结构进行初始化，包括设定输入层神经元个数M、隐藏层神经元个数L和输出层神经元个数N；

[0175] 模拟退火遗传算法优化模块，用于优化所述反向传播神经网络，包括对该训练数据进行长度为K的实数编码，其中，

[0176] K=权值个数+偏倚个数=(M*L+L*N)+(L+N)，

[0177] 则每个编码后的该训练数据包含该反向传播神经网络的权值和偏倚信息，由适应度函数计算编码后的该训练数据的适应度值所确定的概率，找到最具适应度训练数据；

[0178] 反向传播神经网络模型训练模块，用于训练网络安全态势的预测模型，包括将该最具适应度训练数据中的该输入层神经元个数M所对应的安全态势值作为输入值，该输出层神经元个数N所对应的安全态势值作为期望输出值，根据向前传递该输入值、向后传播该期望输出值来训练该反向传播神经网络，从而建立网络安全态势的预测模型；

[0179] 模型预测模块，用于将该输入层神经元个数M所对应的安全态势值作为输入值，根据该网络安全态势的预测模型对该输出层神经元个数N所对应的网络安全态势值进行预测。

[0180] 本发明所提出的网络安全态势预测系统，其中，所述模拟退火遗传算法优化模块进一步包括以下模块：

[0181] 初始值设定模块，用于设定初始温度T_0、最小温度T_{min}、温度T的迭代次数c、训练数据个数S、适应度阈值F和进化代数G，并对所述训练数据进行长度为K的实数编码；

[0182] 适应度值计算模块，用于计算编码后的训练数据的适应度值，适应度Fitness的计算公式为，

[0183] $$Fitness=\alpha(\sum_{i=1}^{n}abs(T_i-O_i)),$$

[0184] 其中，α为函数系数，n为输出层神经元个数，T_i为输出层第i个神 经元的安全态势值的期望输出值，O_i为输出层第i个神经元的安全态势值的预测输出值；

[0185] 选择操作模块，用于选择操作使用适应度比例选择方法，选择出适应度不小于适应度阈值F的训练数据；

[0186] 交叉操作模块，用于根据交叉率交换训练数据的某些基因，将有益基因组合在一起；

[0187] 变异操作模块，用于对训练数据的某些基因座上的基因值作变动，以维持该训练

数据的多样性；

[0188] Metropolis准则接受模块，用于对经过变异步骤后的训练数据按照Metropolis准则进行接受；

[0189] 终止条件判断模块，用于判断是否满足终止条件，若满足终止条件，则直接进入下述最具适应度的训练数据选择模块，若不满足终止条件，则进入下一步骤，其中，终止条件为到达所述训练数据预设定的最大进化次数，或连续多个新解未被接受，或达到预设最低温度T_{min}；

[0190] 迭代次数更新模块，当T_0向着T_{min}的方向逐渐降温时，用于更新迭代次数，并转至所述适应度值计算模块，进行下一轮迭代；

[0191] 最具适应度的训练数据选择模块，用于选择适应度最大的个体作为最具适应度的训练数据。

[0192] 本发明所提出的网络安全态势预测系统，其中，所述Metropolis准则接受模块进一步包括以下模块：

[0193] 能量变化值计算模块，用于计算所述训练数据的进化代数G_t能量变化值$\Delta E=E(G_t)-E(G_{t-1})$，其中$E(G)$为能量的评价函数，取所述适应度值计算模块中所述的适应度为能量评价函数；

[0194] 接受模块，用于接受G_t作为新的训练数据，若$\Delta E<0$则接受G_t作为新的训练数据，若$\Delta E>0$则以概率$p=e^{\frac{-\Delta E}{KT}}$接受$G_t$作为新的训练数据。

[0195] 本发明所提出的网络安全态势预测系统，其特征在于，所述反向传播神经网络模型训练模块进一步包括以下模块：

[0196] 权值和偏倚的初始化赋值模块，用于利用所述模拟退火遗传算法优化模块得到的所述最具适应度的训练数据对应的权值和偏倚，对神经网络的权值和偏倚进行初始化赋值；

[0197] 输入前向传递模块，该模块包括用于判断所述反向传播神经网络训练次数是否满足迭代次数的训练次数判断模块以及隐藏层输出计算模块和输出层输出计算模块；

[0198] 输出后向传播模块，该模块包括用于计算输出层输出误差的输出层误差计算模块、用于判断输出层输出误差是否小于预设定阈值的输出层误差判断模块、用于计算隐藏层误差的隐藏层误差计算模块、用于更新学习速率的学习速率更新模块以及权值更新计算模块和偏倚更新计算模块；

[0199] 网络安全态势的预测模型确立模块，用于根据所述权值和偏倚的初始化赋值模块、所述输入前向传递模块和所述输出后向传播模块的计算与判断后，确立最终的网络安全态势的预测模型。

[0200] 本发明所提出的网络安全态势预测系统，其特征在于，所述输入前向传递模块进一步包括，

[0201] 训练次数判断模块，用于判断所述神经网络的训练次数是否满足迭代次数，若满足迭代次数，则可确立预测模型，若不满足迭代次数，则进入下一步骤；

[0202] 隐藏层输出计算模块，用于计算隐藏层输出，具体为，输入的安全态势值向量$<ns_1,ns_2,\cdots,ns_M>$通过输入层不发生任何变化，即对于输入单元，它的输出O_j等于它的输入

值ns_j，到达隐藏层后，隐藏层的净输入用其输入的线性组合计算$I_j=\Sigma_i w_{ij}0_i+\theta_j$，其中，j=1，2，…h，h为隐藏层神经元个数，$w_{ij}$是由上一层的神经元i到神经元j连接的权重值，$0_i$是i的输出，$\theta_j$是j的偏倚，由神经元激励函数func计算得到隐藏层神经元j的输出0_j，$0_j=func(I_j)$，j=1，2，…h；

[0203] 输出层输出计算模块，用于计算输出层输出，具体为，根据隐藏层输出0_j，计算输出层输出0_k，

[0204] $0_k=\Sigma_j w_{jk}0_j+\theta_k$，

[0205] 其中，k=1，2，…n，n为输出层神经元个数，w_{jk}是由上一层的神经元j到神经元k的连接的权重值，0_j是j的输出，θ_k是k的偏倚；

[0206] 本发明所提出的网络安全态势预测系统，其特征在于，所述输出后向传播模块进一步包括，

[0207] 输出层误差计算模块，用于计算输出层输出误差，具体为，对于输出层单元k误差Err_k的计算公式如下，

[0208] $Err_k=0_k(1-0_k)(T_k-0_k)$，

[0209] 其中，T_k为期望输出的目标值，即真实获得的安全态势值ns_k，0_k为单元k输出的预测态势值ns_k'；

[0210] 输出层误差判断模块，用于判断输出层输出误差Err_k是否小于预设定阈值，若小于预设定阈值，则可确立预测模型，若不小于预设定阈值，则进入下一步骤；

[0211] 隐藏层误差计算模块，用于计算隐藏层误差，下一个较高层隐藏层单元j的误差Err_j的计算公式为如下，

[0212] $Err_j=0_j(1-0_j)\Sigma_k Err_k w_{jk}$；

[0213] 学习速率更新模块，用于更新学习速率，假设可以利用上一轮t-1误差err_{t-1}和这一轮t误差err_t的增大减小的变化对lr_t进行微调，则lr_t的学习速率公式如下，

[0214] 当$err_t>err_{t-1}$时，$lr_t = lr_{t-1} - lr_{t-1} \times \frac{t_{max}-t}{t_{max}} \times \left|\frac{err_t-err_{t-1}}{err_{t-1}}\right|$，

[0215] 当$err_t<err_{t-1}$时，$lr_t = lr_{t-1} + lr_{t-1} \times \frac{t_{max}-t}{t_{max}} \times \left|\frac{err_t-err_{t-1}}{err_{t-1}}\right|$，

[0216] 其中，t_{max}为预设定的最大迭代次数，t为当前进行的迭代轮数；

[0217] 权值更新计算模块，用于计算权值更新，其更新公式如下，

[0218] $w_{ij}=w_{ij}+lr_t Err_j 0_i$，$w_{jk}=w_{jk}+lr_t Err_k 0_j$，

[0219] 其中，lr_t为该轮迭代的学习速率；

[0220] 偏倚更新计算模块，用于计算偏倚更新，输出层θ_k、隐藏层θ_j的更新公式如下，

[0221] $\theta_k=\theta_k+lrErr_k$，$\theta_j=\theta_j+lrErr_j$；

[0222] 最后，进入所述训练次数判断模块，重复进行下一个周期。

[0223] 具体实施例中将采集网络态势值的时间间隔设定为5分钟，即每隔5分钟产生一个当前时刻的网络安全态势值，设定输入层神经元个数M=6、隐藏层神经元个数L=7和输出层神经元个数N=1，即将历史数据时间粒度界定为30分钟，利用上述历史数据预测下一时刻（下一个5分钟）的态势值；设定初始温度T_0=100，最小温度T_{min}=0，温度T的迭 代次数c=50；种群个数S=10，进化代数G=15，初始学习速率lr=0.01，有限次迭代L=200，误差阈值e=0.02，

同时技术方案中相关步骤的其他公式及参数的设定值，如下所示，

[0224] 在步骤33的选择操作中，个体i的选择概率p_i为

[0225] $$p_i = \frac{Fitness_i}{\sum_{j=1}^{S} Fitness_j}$$

[0226] 其中，$Fitness_i$为个体i的适应度值，S为种群个体数目。

[0227] 在步骤34交叉操作中，第s个个体a_s和第l个个体a_l在第j位置交叉如下：

[0228] $$\begin{cases} a_{sj} = a_{sj}(1-b) + a_{lj}b \\ a_{lj} = a_{lj}(1-b) + a_{sj}b \end{cases}$$ 其中，b=0.3。

[0229] 在步骤35的变异操作中，第i个个体的第j位置a_{ij}进行变异如下：

[0230] $$a_{ij} = \begin{cases} a_{ij} + (a_{ij} - a_{max}) * f(G) & \gamma_1 \geq 0.5 \\ a_{ij} + (a_{min} - a_{ij}) * f(G) & \gamma_1 < 0.5 \end{cases}$$

[0231] 其中，γ_1为[0,1]间的随机数，a_{min}为权值或偏倚a_{ij}的下界，a_{max}为权值或偏倚a_{ij}的上界，$f(G) = \gamma_2(1 - \frac{G}{G_{max}})$，$\gamma_2$=0.1，G为当前进化代数，$G_{max}$是预设定的最大进化代数；

[0232] 在步骤422的隐藏层输出计算中，神经元激励函数$func(I_j) = \frac{1}{1+e^{-I_j}}$，j=1,2,…h

[0233] 图3为本发明的技术方案的具体流程说明图，如图中所示，应用模拟退火遗传算法优化反向传播神经网络的网络安全态势预测方法，首先将时间间隔5分钟获得的网络安全态势值作为历史数据集合，由此来准备预测模型的训练数据；接下来，利用训练数据中的前6个时刻的态势值为输入，以后1个时刻的态势值为期望输出，训练所述的模拟退火遗传算法优化的反向传播神经网络，当满足最大训练迭代次数200或者满足误差阈值0.02时，预测模型学习训练完成；最终利用包括当前时刻在内的前6个时刻的态势值为输入，预测未来下1个时刻的态势值。如图4所示，其中节点为方形的线代表每个时刻的真实态势值，节点为棱形的线代表每个时刻的预测态势值，节点为三角形的线代表该时刻预测 态势值相对于真实态势值的误差。

[0234] 在该网络安全态势预测方法中，一次性学习完成的神经网络对于后续数据的预测能力会随着时间的推移而逐渐退化，为满足网络态势值随时间变化和系统应用的实际情况，每来一组或者几组短时间内的态势值向量，就应当进入网络学习，计算误差，调整神经元连接权值和偏倚，来达到“再适应”实时态势值变化的要求，同时，采用多个依据本发明训练的网络结构预测态势值加权取平均值，以T为时间周期和预测误差阈值e来训练和更新上述网络结构。态势值训练数据集合分为历史数据、近期数据和预测数据三个部分，历史数据是由据当前时间距离较远的大量的充分的真实态势值组成，用于神经网络的初次训练和学习；近期数据是据当前时间较近的真实的态势值，用于验证训练模型的有效性，以及对未来时刻态势值的预测作为输入向量；预测数据是预测过程的输出结果。当到达训练时间周期T或预测态势值与真实态势值误差达到或超过预设定的误差阈值e，则取近期数据中预测态势值和对应的真实态势值间的误差来修正连接权值、偏倚以及学习速率，修改方式参照步

骤434、步骤435和步骤436。

[0235] 最后应当说明的是，本发明并非限制于这里所描述的实施例，任何对本发明的技术方案的修改或者等同替换，都不脱离本发明技术方案的范围，均应涵盖在本范围的权利要求范围内。

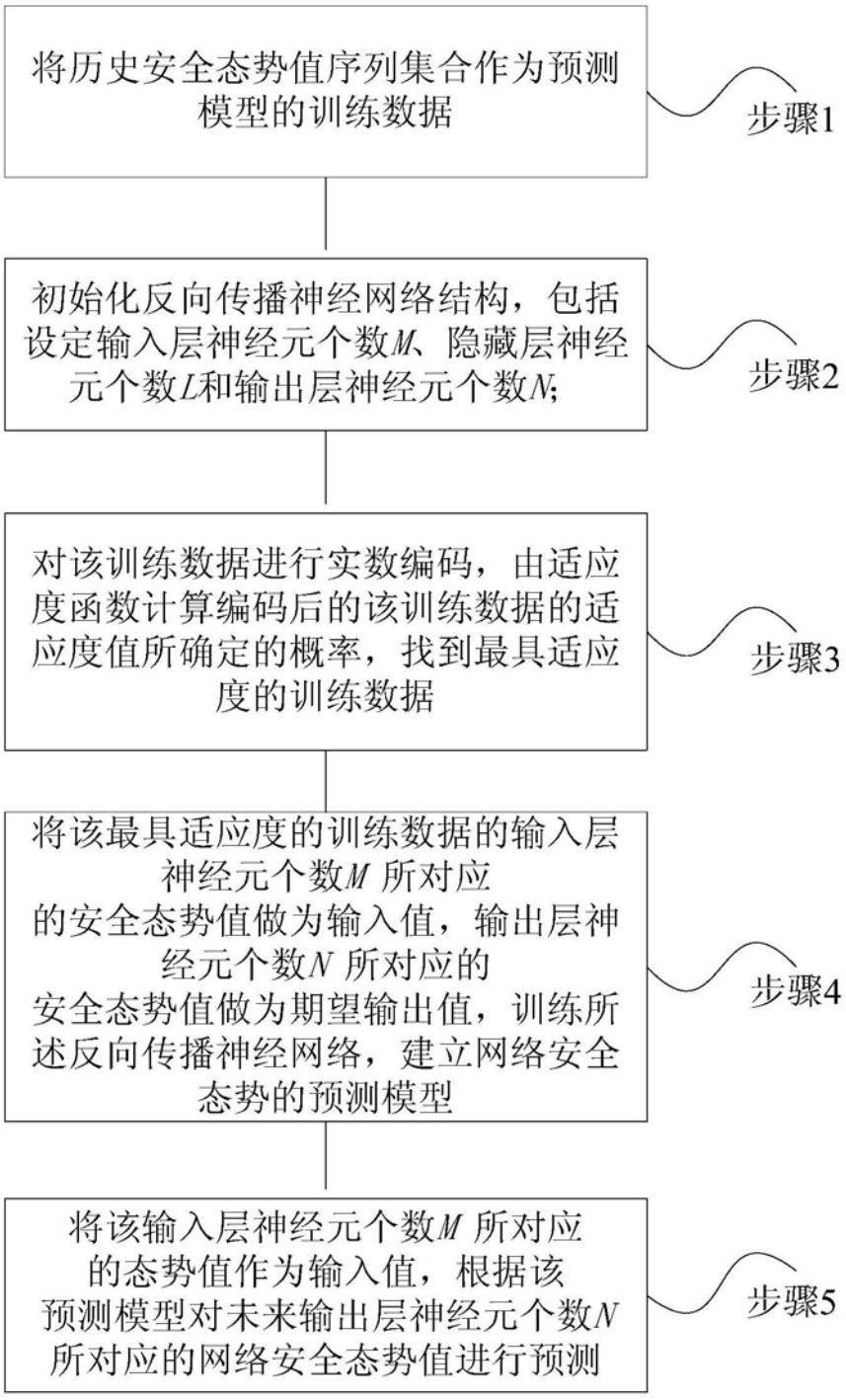
将历史安全态势值序列集合作为预测模型的训练数据
步骤1
初始化反向传播神经网络结构，包括设定输入层神经元个数M、隐藏层神经元个数L和输出层神经元个数N;
步骤2
对该训练数据进行实数编码，由适应度函数计算编码后的该训练数据的适应度值所确定的概率，找到最具适应度的训练数据
步骤3
将该最具适应度的训练数据的输入层神经元个数M 所对应的安全态势值做为输入值，输出层神经元个数N 所对应的安全态势值做为期望输出值，训练所述反向传播神经网络，建立网络安全态势的预测模型
步骤4
将该输入层神经元个数M 所对应的态势值作为输入值，根据该预测模型对未来输出层神经元个数N 所对应的网络安全态势值进行预测
步骤5

图1

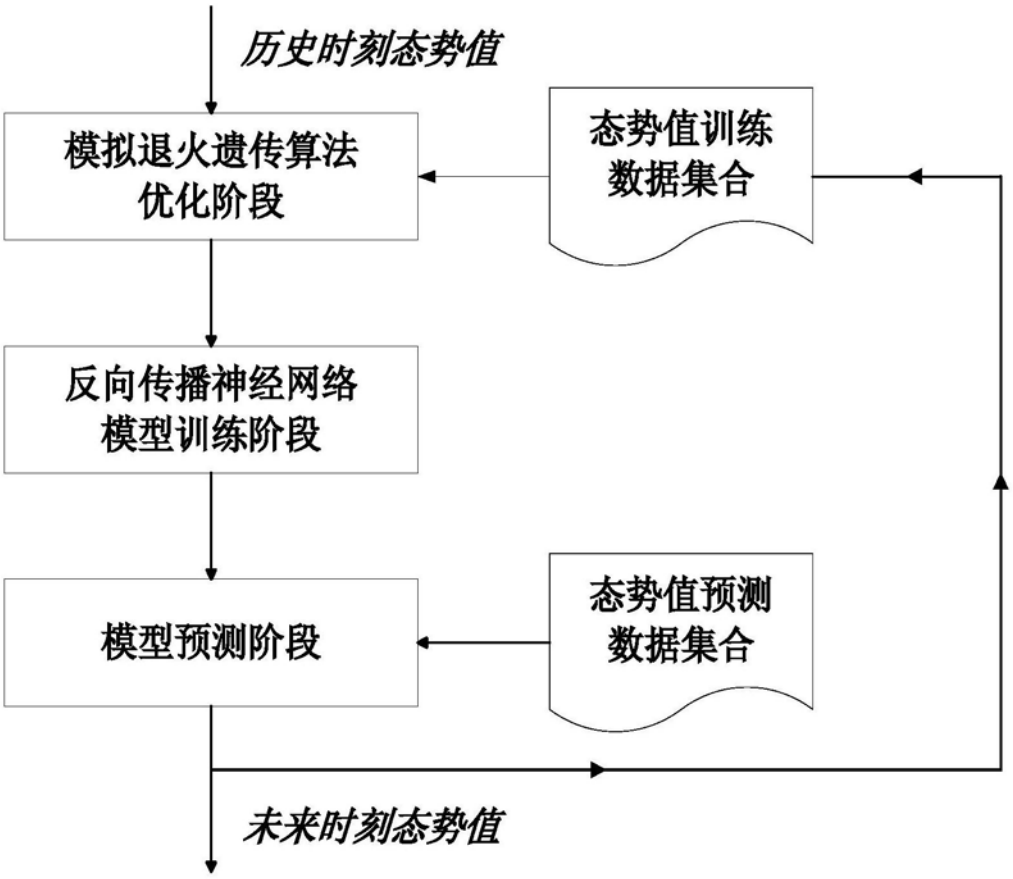
历史时刻态势值
模拟退火遗传算法
优化阶段
态势值训练
数据集合
反向传播神经网络
模型训练阶段
模型预测阶段
态势值预测
数据集合
未来时刻态势值

图2

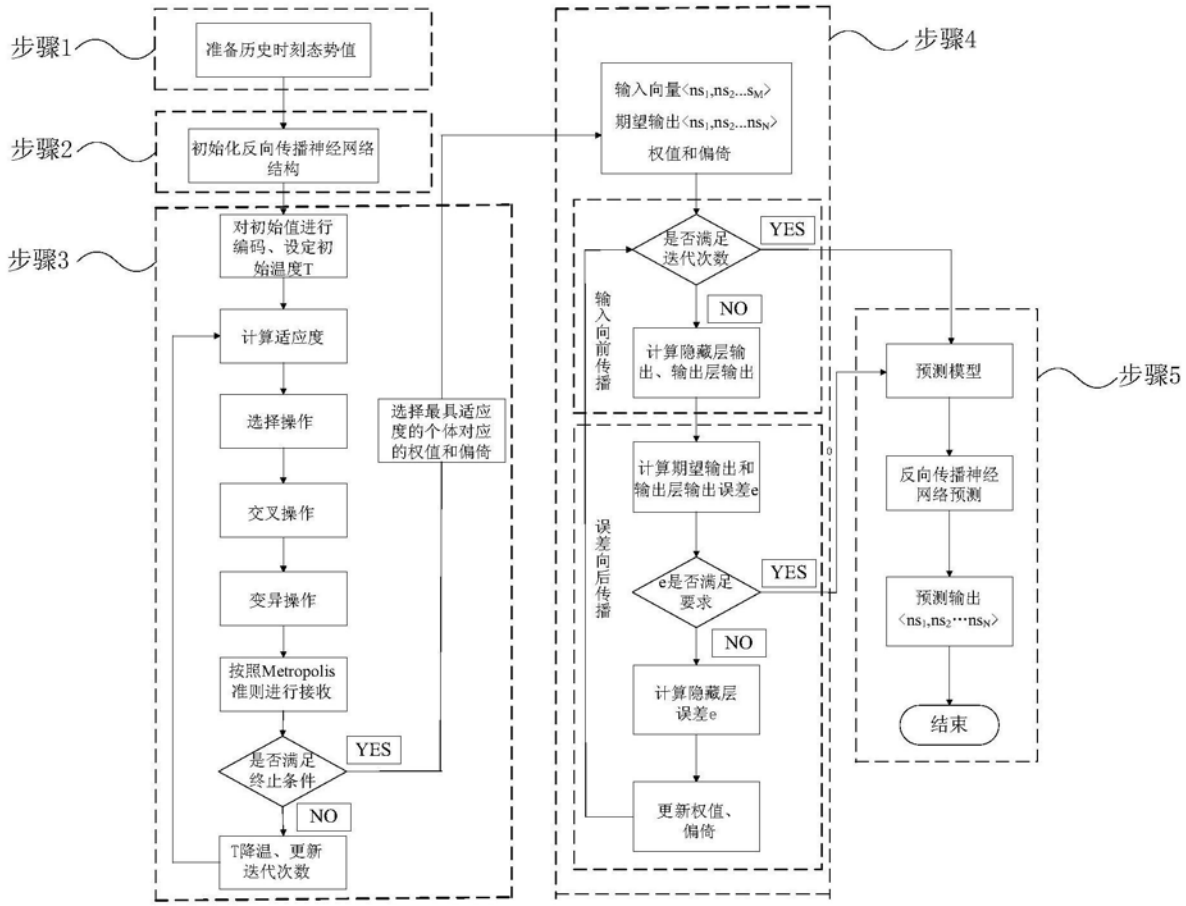
步骤1
准备历史时刻态势值
步骤2
初始化反向传播神经网络结构
步骤3
对初始值进行编码、设定初始温度T
计算适应度
选择操作
交叉操作
变异操作
按照Metropolis准则进行接收
是否满足终止条件
YES
NO
T降温、更新迭代次数
选择最具适应度的个体对应的权值和偏倚
步骤4
输入向量<ns1,ns2...sM>
期望输出<ns1,ns2...nsN>
权值和偏倚
是否满足迭代次数
YES
NO
输入向前传播
计算隐藏层输出、输出层输出
计算期望输出和输出层输出误差e
误差向后传播
e是否满足要求
YES
NO
计算隐藏层误差e
更新权值、偏倚
步骤5
预测模型
反向传播神经网络预测
预测输出<ns1,ns2…nsN>
结束

图3

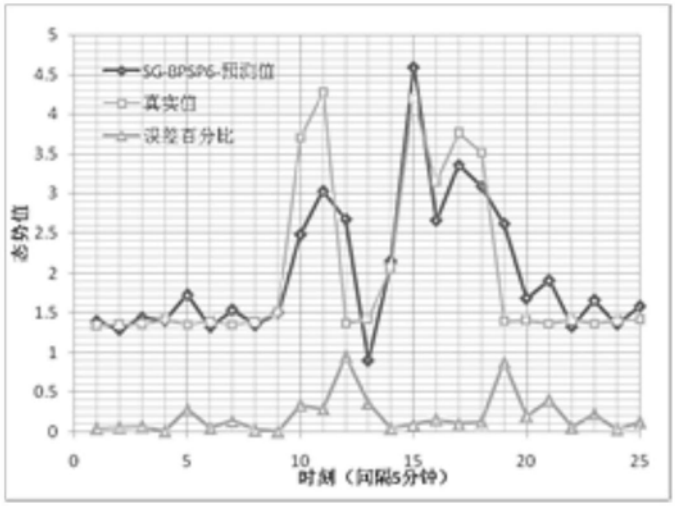
SG-BPSPS-预测值
真实值
误差百分比
态势值
时刻（间隔5分钟）
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
5
0
5
10
15
20
25

图4