2020 年 10 月 第 10 期
Vol. 41 No. 10 2020

一种基于 SA-SOA-BP 神经网络的网络安全态势预测算法

张 然,刘 敏,张启坤,尹毅峰

(郑州轻工业大学 计算机与通信工程学院,郑州 450002)
E-mail: ranranzh@ sina. com

摘 要: 网络安全态势预测能够依据已有的网络安全数据预测网络未来的安全状况及其变化趋势,为安全策略的选取提供指导,从而增强网络防御的主动性,尽可能地降低危害. 然而现有的网络安全态势预测方法的精准度和收敛性还不理想. 为了提高网络安全态势预测的准确性,提出了一种将模拟退火算法(SA)引入人群搜索算法(SOA)优化 BP 神经网络的网络安全态势预测方法. 该算法利用人群搜索算法特有的利己、利他、预动和不确定推理四大行为特征确定搜索策略,找到最佳适应度个体,获取最优权值和阈值,然后再对 BP 神经网络的随机初始阈值和权值进行赋值,经过神经网络训练后得到预测值. 针对人群搜索算法在搜索后期易陷入局部最优和收敛缓慢等问题,又将模拟退火算法引入人群搜索算法,根据它的 Metropolis 准则以一定的概率接受恶解,避免了算法陷入局部最优的陷阱,提高了该算法的全局搜索能力. 与其它基于改进 BP 神经网络的预测算法进行对比的实验表明,该优化算法准确性更高,稳定性更强,收敛效果更好.

关 键 词: BP 神经网络; 人群搜索算法; 模拟退火算法; 网络安全; 态势预测
中图分类号: TP393 **文献标识码:** A **文 章 编 号:** 1000-1220(2020)10-2157-07

Network Security Situation Prediction Algorithm Based on SA-SOA-BP Neural Network

ZHANG Ran,LIU Min,ZHANG Qi-kun,YIN Yi-feng
(School of Computer and Communication Engineering,Zhengzhou University of Light Industry,Zhengzhou 450002,China)

Abstract: Network security situation prediction can predict the future security situation and its changing trend based on the existing network security data,and provide guidance for the selection of security strategies,so as to enhance the initiative of network defense and reduce the harm as much as possible. However,the accuracy and convergence of existing network security situation prediction methods are not ideal. In order to improve the accuracy and convergence of network security situation prediction,a network security situation prediction method based on improved BP neural network optimized by introducing simulated annealing(SA) algorithm into seeker optimization algorithm(SOA) is proposed. This algorithm uses the four behavioral characteristics of seeker optimization algorithm: self-interest,altruism,pre-action and uncertain reasoning to determine the search strategy,find the best fitness individual,obtain the best weight and threshold value,then assign them to the random initial threshold and weight value of BP neural network,and get the prediction value after the training of neural network. But the seeker optimization algorithm is prone to fall into local optimization and slow convergence in the later stage of the search. In order to improve its deficiencies,the simulated annealing algorithm is introduced into the seeker optimization algorithm. According to the Metropolis criterion of the simulated annealing algorithm,the bad solution is accepted with a certain probability,which avoids the algorithm falling into the trap of local optimization and improves the global search ability of the algorithm. Compared with other prediction algorithms based on optimized BP neural network,the experimental results show that this optimized algorithm has higher accuracy,stronger stability and better convergence.

Key words: BP neural network; seeker optimization algorithm; simulated annealing algorithm; network security; situation prediction

1 引 言

随着大数据、人工智能和互联网的高速发展及应用,网络结构的复杂化、数据的多元化以及网络协议的多样化,使得多层面、多形式的网络安全风险随之加剧. 网络攻击方式也变得多样化,日渐向着分布化、规模化、复杂化等方向发展,而像入侵检测系统、防火墙等传统的网络安全防御手段已经不能满足现在高速、智能、多源的网络安全需求,我们需要更加先进、优化的技术手段和方式方法去防范网络安全事件的发生.

安全态势感知最早运用在航空和军事领域,用来快速决策和处理复杂的航空和军事事件. 1999 年,Bass 首次提出了网络安全态势感知(network security situation awareness,NSSA)的概念[1],后来态势感知被研究者们广泛地运用到网络安全领域. 研究者们发现态势感知不仅可以进行网络安全态势评估,也可以用来进行网络安全态势预测,使原来的被动防御变为主动防御,可以很大程度的解决网络安全的防御问题,

收稿日期: 2020-04-14 收修改稿日期: 2020-06-01 基金项目: 河南省重点科技攻关项目(142102210081)资助; 国家自然科学基金项目(61772477)资助; 河南省产学研合作项目(132107000066)资助. 作者简介: 张 然,女,1973 年生,博士,副教授,CCF 会员,研究方向为网络信息安全、入侵检测; 刘 敏,女,1995 年生,硕士研究生,CCF 会员,研究方向为网络信息安全、机器学习; 张启坤,男,1980 年生,博士,副教授,研究方向为信息安全与密码学; 尹毅峰,男,1971 年生,博士,教授,CCF 会员,研究方向为信息安全与密码学.

因此成为了当下的一个热点研究方向.

本文的主要工作如下:

1) 提出了一种基于 SA-SOA-BP 神经网络的网络安全态势预测方法,利用改进的 BP 神经网络模型训练产生安全态势值来预测未来网络安全态势的状况以及发展趋势.

2) 将人群搜索算法应用到 BP 神经网络中优化它的权值和阈值,提高基于神经网络的网络安全态势预测的精准度.

3) 将模拟退火算法引入到人群搜索算法中来克服它的局部最优和收敛缓慢的问题,进一步提高安全态势预测的精准度和收敛速度.

4) 应用不同优化算法进行实验对比,表明我们提出的基于 SA-SOA-BP 神经网络的网络安全态势预测方法的误差最小,精准度最高,收敛速度较快,可以有效预测未来网络安全态势的变化情况.

2 相关工作

目前,网络安全态势预测的研究已经非常广泛,国外对这方面的研究相对比较早. 文献 [2] 将现有的网络安全状况预测机制分为三大类,并就每种模型的优缺点进行了回顾. 文献 [3] 提供了有关网络安全中的预测和方法的概述,对攻击预测、意图识别、入侵预测和网络安全态势预测进行了讨论和比较. 文献 [4] 提出了一种基于变长马尔可夫的预测模型,通过捕获攻击轨迹的顺序属性,实现对该攻击的预测. 文献 [5] 提出了一种基于时间序列波动分析和预测的方法,实现对分布式拒绝服务(DDoS) 活动的预测. 文献 [6] 提出了基于语义 Web 的网络安全状态预测工具,它可以在系统配置不断变化的领域(如计算机网络) 中应用.

近些年国内在网络安全态势预测方面也做了大量的研究,很多研究者们尝试将人工智能的方法引入态势感知领域来提高态势评估和预测的准确性. 文献 [7] 提出了一种基于灰色关联分析和支持向量机的网络安全态势预测方法,该方法采用灰色关联分析法(GRA) 对网络评估指标进行加权分析,利用支持向量机(SVM) 算法对预测过程进行仿真,提高了预测的精准度. 文献 [8] 提出了一种基于改进的深度神经网络模型的入侵检测方法,该方法利用自编码器对数据特征学习、降维和去冗,利用深度神经网络分类,再通过多层网格搜索算法优化,很大程度上提高了精准度和训练速度. 文献 [9] 提出了一种基于时空维度分析的网络安全态势预测方法,解决了未来安全态势要素值的新变化以及周围节点安全态势要素对态势预测的影响,但此方法受提取的态势要素值的影响较大,另外该方法中用到了脆弱性预测算法,它的随机性较大,对预测的准确性有影响. 文献 [10] 提出了一种基于隐 Markov 模型的实时网络安全态势预测模型,可以有效的提高预测的实时性,但是 HMM 模型存在参数估计的效率性问题. 文献 [11] 提出了一种基于优化 RBF 神经网络的 SDN 网络安全态势评估方法,该方法在 SDN 网络中能全面提取态势指标,再通过改进的 K-means 和 PSO 算法优化 RBF 神经网络,在一定程度上提高了态势评估的准确性和性能,但在提取态势指标时还存在不足. 文献 [12] 提出了一种基于模拟退火算法和变步长学习策略优化 BP 神经网络的评估模型,克服了传统 BP 神经网络反馈误差慢和易陷入局部极值的局限性,提高了评估的精准度. 文献 [13] 提出了一种基于 MapReduce 和 SVM 的网络安全态势预测模型,该模型使用杜鹃算法优化 SVM 的参数,并使用 MapReduce 对 SVM 进行分布式训练,以提高训练速度. 文献 [14] 提出了一种基于 IFS-NARX 模型的网络安全态势预测方法,该方法具有较高的学习效率,可以更及时、准确地预测网络安全态势. 文献 [15] 提出了一种基于灰色神经网络的云环境中网络安全态势预测方法,该方法用于解决云环境中现有的网络安全态势预测在准确性和实时性能方面的局限性. 这些预测模型和方法是近几年研究中比较常用的,主要有基于向量机模型、时间序列分析模型、要素分析方法、隐 Markov 模型、神经网络模型等,这些改进的方法相比传统方法在一定程度上提高了评估预测的效果,但是这些方法面对收集到的海量数据,在准确性和效率上还不理想,无法适应动态多变的网络安全需求.

人群搜索算法(SOA) 是一种智能搜索算法,应用较为广泛. 在前面的研究中我们采用 SOA 优化 BP 神经网络并应用于网络安全态势预测,从一定程度上提高了预测的精准度,但 SOA 存在搜索后期易陷入局部最优即“早熟”现象和收敛缓慢的问题. 为了克服这个问题,本文又将模拟退火算法(SA) 引入人群搜索算法(SOA) 并与 BP 神经网络相结合,对基于 SA-SOA-BP 神经网络的网络安全态势预测算法进行研究和仿真实验.

3 基于 SA-SOA-BP 神经网络的网络安全态势预测

BP(Back Propagation) 神经网络是 1986 年由 Rumellhart 和 McCelland 为首的科学小组提出[16]. 由于 BP 神经网络结构简单,可调整的参数多,训练的算法多,鲁棒性和自我学习能力强,并且可操作性好,因此它是目前应用最广泛的神经网络预测模型之一[17]. BP 神经网络主要是通过反向传播算法反复调整网络的权值和阈值,直到得到最优的权值和阈值,再经过不断地学习和训练,使输出数据与真实值尽量地一致,最后当输出的误差平方和小于指定的误差时,训练完成,保存最优连接权值和阈值. 但是它的初始连接权值和阈值难以准确获得并且迭代次数多、运算速度低,不能保证收敛到全局极值点. 针对 BP 神经网络的这些局限性,大多数的研究都是通过智能优化算法寻找最优权值和阈值来弥补 BP 神经网络的不足,比如采用粒子群算法(PSO) 优化 BP 神经网络[18],采用遗传算法(GA) 优化 BP 神经网络[19] 等.

为了弥补 BP 神经网络的局限性,本文将人群搜索算法(SOA) 应用到 BP 神经网络中,迭代寻找其最优权值和阈值. 但是 SOA 算法在寻找最优个体的过程中易陷入局部最优和出现收敛缓慢等问题,因此又将模拟退火算法(SA) 引入到人群搜索算法(SOA) 中,提高它的全局搜索能力,并将模拟退火算法优化的人群搜索算法(SA-SOA) 与 BP 神经网络相结合进行网络安全态势预测,以此提高基于 BP 神经网络进行网络安全态势预测的效率和准确性.

3.1 人群搜索算法

人群搜索算法(Seeker Optimization Algorithm,简称 SOA) 是一种较新的启发式随机搜索算法[20]. 它主要研究和

分析人类在随机搜寻过程中的智能行为,依靠人类的社会经验,同时结合进化的思想,以搜索最优位置为核心,通过利己、利他、预动和不确定这四种搜索策略行为对其进行建模,确定人群搜索的方向和步长,然后不断地更新位置,获得最优解. SOA 算法的优点在于它简单,概念明确、清晰、易于理解、收敛速度快、收敛精度高. 其计算步骤主要包括:

1) 搜索步长的确定

在确定步长时,需要对个体最优适应度值进行降序排列,并给每个个体赋予索引号作为模糊推理的输入,本文是使用高斯线性隶属函数来表示搜索步长的模糊变量的输出,它可以很好的将第 i 个个体最佳适应度值线性的映射到最小和最大隶属度之间,映射公式如下:

$$u_i = U_{max} - (sizepop - Indexfitnessgbest(i)) * \frac{U_{max} - U_{min}}{sizepop - 1} \tag{1}$$

$$u_{ij} = u_i + (1 - u_i) * rand(j = 1,2,3,\cdots,D) \tag{2}$$

上式中,u_i 为第 i 个个体对应的隶属度; $Indexfitnessgbest(i)$ 为第 i 个个体最佳适应度值的索引号; $sizepop = 30$, $sizepop$ 为种群规模; $U_{max} = 0.95$ 和 $U_{min} = 0.0111$ 分别表示最大和最小函数隶属度; u_{ij} 表示在 j 维探索空间中目标函数值 i 所对应的隶属度; 根据公式(1)和公式(2)得到最佳适应度值个体对应的隶属度,再根据公式(3)确定步长:

$$a_{ij} = \delta_{ij}\sqrt{-log(u_{ij})} \tag{3}$$

上式中,a_{ij} 表示为第 i 个搜寻者在 j 维搜索空间的搜索步长,δ_{ij} 为高斯隶属函数的参数,其值由下列公式确定.

$$\delta_{ij} = H(t) * |zbest - 5 * rands(1,10)| \tag{4}$$

$$H(t) = \frac{maxgen - t}{maxgen} \tag{5}$$

$zbest$ 表示为全局最佳; $rands(1,10)$ 表示为 [1,10] 之间的随机实数; $H(t)$ 表示第 t 次迭代的权重函数值,在迭代的过程中是不断的变化,它受最大迭代次数和当前迭代次数的影响,其中 $maxgen = 100$.

2) 搜索方向的确定

在确定搜索方向时,依据个体最佳和全体最佳与当前个体相比较确定搜索方向是利己、利他还是预动方向.

$$\vec{d}_{i,ego}(t) = \vec{g}_{i,best} - \vec{x}_i(t) \tag{6}$$

$$\vec{d}_{i,all}(t) = \vec{z}_{i,best} - \vec{x}_i(t) \tag{7}$$

$$\vec{d}_{i,pro}(t) = \begin{cases} \vec{D}_i & F_{i,best} < F_{x_i} \\ -\vec{D}_i & F_{i,best} \geqslant F_{x_i} \end{cases} \tag{8}$$

上式中,$\vec{x}_i(t)$ 表示第 i 个个体在第 t 次迭代时的最佳位置,$\vec{g}_{i,best}$ 表示第 i 个个体到当前为止经历的最佳位置,$\vec{z}_{i,best}$ 表示为第 i 个个体所在邻域的集体历史最佳位置,$F_{i,best}$ 表示为 $\vec{g}_{i,best}$ 位置的适应度值,F_{x_i} 表示为 $\vec{x}_i(t)$ 位置的适应度值. 利己方向的确定是以个体目前最佳位置与个体在第 t 次迭时的最佳位置相减为确定标准,利他方向的确定是以全局最佳位置与个体在第 t 次迭时的最佳位置为确定标准,而预动方向的确定是以个体当前适应度值与个体最佳适应度值相比较为确定标准.

基于 SA-SOA-BP 神经网络的网络安全态势预测算法是以三个方向的随机加权几何平均数为标准确定搜索方向,其计算公式如下:

$$\vec{d}_{ij}(t) = sign(W\vec{d}_{ij,pro} + \varphi_1\vec{d}_{ij,ego} + \varphi_2\vec{d}_{ij,alt}) \tag{9}$$

$$W = W_{max} - t * \frac{W_{max} - W_{min}}{maxgen} \tag{10}$$

在上式中,W 为惯性权值,φ_1 和 φ_2 为 [0,1] 内均匀分布的常数. t 为当前迭代次数,取值范围为 [2,$maxgen$] 之间的整数,$W_{max} = 0.9$ 为权重最大值,$W_{min} = 0.1$ 为权重最小值.

3) 位置更新

计算得到个体探索的方向和步长之后,要对个体的位置进行更新. 位置更新公式如下:

$$x_{ij}(t+1) = x_{ij}(t) + a_{ij}(t) * \vec{d}_{ij}(t) \tag{11}$$

人群搜索算法属于一种优化算法,一般都是通过此算法去优化另一种算法或模型,使优化后的算法或模型具有更好的准确性、稳定性、收敛性和有效性等. 它为一些传统的预测模型提供了帮助.

3.2　模拟退火算法

模拟退火算法(Simulated Annealing,简称 SA)最早是由 Metropolis [20] 在 1953 年提出,并由 Kirkpatrick [22] 等成功引入组合优化领域. 其思想是对高温固体退火降温过程进行模拟,通过加温、等温和冷却这三个过程,将系统的能量看成优化问题的目标函数,其能量随着温度的降低也随之下降,当温度缓缓降温并趋于零时,此时既是能量最低状态,也是得到相对全局最优解的时刻 [23]. SA 算法具有较强的鲁棒性、隐含并行性、广泛的适应性,以及全局搜索能力,它主要利用 Metropolis 算法并适当控制温度的下降过程,以一定的概率接受劣质解,跳出陷入局部极值的陷阱,从而提高算法的全局收敛性. SA 算法在大多数情况下也是用于优化另一种算法或模型,很少单独使用.

人群搜索算法(SOA)具有收敛速度快和精度高等优点,但同时也易陷入局部最优,即出现“早熟”现象,原因是人群搜索算法在搜索后期,当搜索步长趋于零时,仍不能搜索到全局最优解. 为了解决这一问题,本文将模拟退火算法(SA)引入到人群搜索算法(SOA)中,SA 算法在搜索寻优过程中会以一定的概率接受劣质解,即 SA 算法既能接受优解又能接受恶解,从此避免了 SOA 算法陷入局部最优.

1) 初始化温度

在退火算法中,初始化温度的确定很关键,它将会直接影响算法的初始性能,若确定不好,最终可能导致无用的搜索和增加搜索时间. 初始温度确定如下:

$$T = -\frac{fitnesszbest}{\log(\alpha)} \tag{12}$$

上式中,$fitnesszbest$ 表示全局最佳适应度值. α 为初始接受概率,一般取值为 [0.2,0.5].

2) 退火速率的控制

$$T_{t+1} = \gamma T_t (2 \leqslant t \leqslant maxgen, 0 \leqslant \gamma \leqslant 1) \tag{13}$$

上式中,γ 为降温的速率,t 为迭代次数.

3) 突跳概率的确定

当退火温度确定时,当前最佳适应度值就以突跳概率代替以前个体最佳适应度值和全局最佳适应度值,则突跳概率的确定公式如下:

$$P=\begin{cases}1 & df<0\\ exp\left(-\frac{df}{T}\right) & df\geqslant 0\end{cases} \tag{14}$$

$$df=fitness(i)-fitnesszbest \tag{15}$$

上式中,$fitness(i)$ 为当前个体适应度值,$fitnesszbest$ 为全局最佳适应度值. 如果 $df<0$,则以概率 1 接受新解;否则以概率 $exp(-df/T)$ 接受新解. 从上式可以看出,在一定程度上退火算法可以帮助人群搜索算法避免陷入局部最优值,最后得到最优解.

3.3 基于 SA-SOA-BP 神经网络的安全态势预测算法

应用模拟退火算法优化人群搜索算法改进 BP 神经网络(简称 SA-SOA-BP 神经网络)的网络安全态势预测方法的主要步骤如下:

Step 1. 预处理样本数据,再根据样本数据的特点确定 BP 神经网络的结构,并初始化 BP 神经网络的连接权值和阈值.

Step 2. 初始化种群个体、种群规模、最大迭代次数、空间维数、最小最大隶属度、权重的最小值最大值.

Step 3. 初始化退火的温度、降温速率、突跳概率.

Step 4. 将预处理过的样本数据代入适应度函数中,计算个体的适应度值,找出全局最佳、个体最佳、个体最佳适应度值和全局最佳适应度值. 此算法是以用训练数据训练 BP 神经网络得到的预测值与真实值之间的误差绝对值和作为个体适应度值.

Step 5. 初始化经验梯度方向、搜索步长和方向以及高斯函数的参数 δ_{ij}. 确定搜索策略,即搜索方向的确定,根据公式(6)-式(7)计算可得;确定经验梯度的方向,根据公式(9)计算可得;确定高斯函数的参数 δ_{ij},根据公式(4)-式(5)计算可得;确定搜索步长的大小,根据公式(3)计算可得;根据计算得到的步长和方向按公式(11)更新位置,更新个体最优和群体最优以及它们的适应度值.

Step 6. Metropolis 准则的引入. 在全局最优适应度值的邻域内选择一个搜寻者,按公式(15)计算当前个体的适应度值与全局最佳适应度值的差值 df. 如果 $df<0$,则以概率 1 接受新的位置,否则以概率 $\exp(-df/T)$ 接受新的位置,然后更新个体最优位置 $gbest$ 和群体最优位置 $zbest$.

Step 7. 降温处理,根据公式(13)控制温度.

Step 8. 判断是否满足循环的终止条件,即是否超过最大迭代次数和种群规模. 如果没有,继续迭代寻优,跳转至 Step 5.

Step 9. 得到最佳网络权值和阈值,并赋值赋给 BP 神经网络的随机初始阈值和权值.

Step 10. 训练及预测. 将处理过的训练数据输入该模型,经过训练,得到具有预测能力的 SA-SOA-BP 模型,再将测试数据输入该模型,得到预测态势值,分析结果.

基于 SA-SOA-BP 神经网络的安全态势预测算法流程如图 1 所示.

4 实验与结果分析

本文将模拟退火算法(SA)引入人群搜索算法(SOA)优化 BP 神经网络应用于网络安全态势预测中来提高预测的准确性和收敛性. 实验以《网络安全信息与动态周报》2015 年第 1 期-2017 年第 7 期所发布的网络安全数据作为实验数据,它主要以感染病毒的主机数量、被篡改的网站总数、被植入后门网站总数、境内网站的仿冒页面数量和新增信息安全漏洞

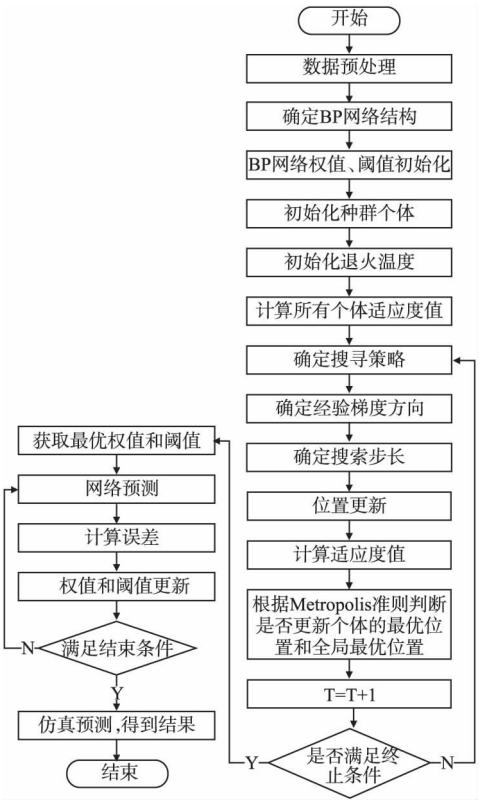

图 1 基于 SA-SOA-BP 的网络安全态势预测算法流程图

Fig. 1 Flow chart of network security situation prediction algorithm based on SA-SOA-BP

数量作为评价指标. 这五方面可以比较全面的反应现代网络安全的状况,可以作为评价每周的网络安全基本态势的指标. 为了实验方便,这里将优、良、中、差、危五个安全等级转化为数字等级,如表 1 所示.

表 1 网络安全态势值转换表

Table 1 Network security situation value conversion table

优	良	中	差	危
5	4	3	2	1

4.1 数据预处理

根据神经网络的特性,训练样本数量过多,将会增加训练的时间,数量过少将会降低预测准确度,所以本文选择 101 条数据为训练样本,10 条数据为测试样本. 为了提高预测的准确度,需要先对数据进行预处理和归一化操作. 数据归一化的方法一般有两种,一种是归一化为 [0,1],一种是归一化为 [-1,1]. 本实验用到了后一种方法,归一化公式见公式(16),归一化后的结果如图 2 所示.

$$y=2\times\frac{x_i-x_{min}}{x_{max}-x_{min}}+(-1) \tag{16}$$

4.2　预测及结果分析

1) 确定 BP 神经网络的网络结构. 由上述所知网络安全态势共有五大评价指标,最后要化为一个安全等级,所以此实验有五个输入参数,一个输出参数,再根据公式(17) -公式(19) 确定隐含层节点个数.

$$l < n - 1 \tag{17}$$

$$l < \sqrt{m + n} + a \tag{18}$$

$$l = log_2 n \tag{19}$$

此公式中,n 为输入层节点数; l 为隐含层节点数; m 为输出层节点数; a 为 0-10 之间的正整数. 再根据试凑法确定此实验的隐含层节点数为 8,最终确定本实验的网络结构为 5-8-1.

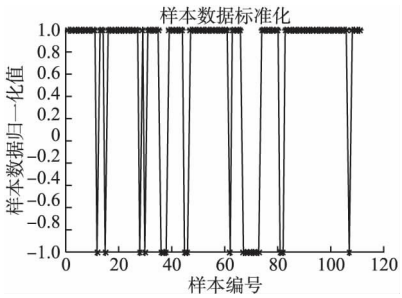

图 2　样本数据标准化

Fig. 2　Standardization of sample data

2) 验证应用 SA-SOA 算法优化 BP 神经网络进行网络安全态势预测的精准性和优越性. 采用均方误差(*MSE*) 、平均绝对百分比误差(*MAPE*) 和均方根误差(*RMSE*) 三个性能指标来衡量真实值与预测值之间的差异.

均方误差指标:

$$MSE = \frac{1}{n}\sum_{t=1}^{n}(x_t - x_t^*)^2 \tag{20}$$

平均绝对百分比误差指标:

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{x_t - x_t^*}{x_t}\right| * 100\% \tag{21}$$

均方根误差指标:

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(x_t - x_t^*)^2} \tag{22}$$

上式中,x_t 和 x_t^* 分别表示真实值和预测值.

3) 与其它优化算法的预测结果进行对比分析. 图 3 显示了基于粒子群算法(PSO) 、遗传算法(GA) 、人群搜索算法(SOA) 和 SA-SOA 算法优化 BP 神经网络进行网络安全态势预测的实验结果对比图,以及各个优化算法的预测值的折线图的变化趋势和真实值的折线图的接近程度.

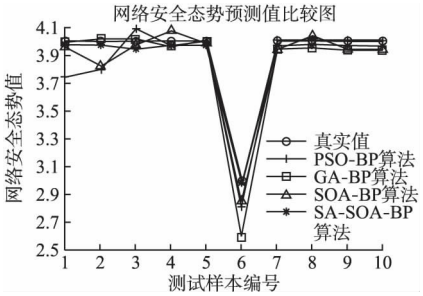

图 3　SA-SOA 与其它算法优化 BP 神经网络的安全态势预测对比图

Fig. 3　Comparison chart of SA-SOA and other algorithms to optimize the BP neural network for security situation prediction

从图 3 可以看到,基于 PSO 算法优化 BP 神经网络得到的预测值折线图在开始部分相对于真实值折线图变化波动比较大,后期相对稳定; 基于 GA 优化 BP 神经网络得到的预测值折线图主要在第 6 个测试样本数据点与真实值的折线图相差较大,其它部分比较接近; 基于 SOA 优化 BP 神经网络得到的预测值折线图相对真实值的折线图变化波动有点大,不太稳定; 而基于 SA-SOA 算法优化 BP 神经网络得到的预测值折线图整体上比较接近真实值折线图,比其它的智能优化算法相对真实值的折线图波动最小,与真实值的折线图更接近,更吻合.

表 2 给出了 PSO、GA、SOA 及 SA-SOA 优化 BP 神经网络进行网络安全态势预测得到的 10 个测试值以及与真实值之间的绝对误差.

表 2　预测数据分析表

Table 2　Prediction data analysis table

真实值	PSO 预测值	PSO 绝对误差	GA 预测值	GA 绝对误差	SOA 预测值	SOA 绝对误差	SA-SOA 预测值	SA-SOA 绝对误差
4	3.742	-0.258	3.994	-0.006	3.958	-0.042	3.976	-0.024
4	3.795	-0.205	4.019	0.019	3.824	-0.176	3.974	-0.026
4	4.094	0.094	4.016	0.016	3.975	-0.025	3.949	-0.051
4	3.966	-0.034	3.959	-0.041	4.075	0.075	3.969	-0.031
4	4.001	0.001	4.010	0.010	3.989	-0.011	3.975	-0.025
3	2.813	-0.187	2.600	-0.4	2.850	-0.150	2.988	-0.012
4	4.012	0.012	3.944	-0.056	3.947	-0.053	3.971	-0.029
4	4.011	0.011	3.957	-0.043	4.041	0.041	3.979	-0.021
4	4.009	0.009	3.938	-0.062	3.946	-0.054	3.970	-0.030
4	4.013	0.013	3.934	-0.066	3.943	-0.057	3.970	-0.030

从表 2 可以看出各个算法在 10 个测试样本数据点得到的预测值以及它们与真实值之间的绝对误差的具体数值,整体上看基于模拟退火算法改进的 SOA-BP 算法的误差更小,这说明基于 SA-SOA-BP 算法进行网络安全态势预测的准确性更高.

表 3 分别计算了 PSO、GA、SOA、SA-SOA 优化 BP 神经网络进行网络安全态势预测得到的预测值与真实值之间的均方误差、平均绝对百分误差以及均方根误差.

从表 3 可以宏观地看出,基于 SA-SOA-BP 神经网络的网络安全态势预测算法得到的预测值与真实值之间的均方误差、平均绝对百分比误差和均方根误差三个衡量指标值相对

表 3 精准度对照表

Table 3 Accuracy comparison table

评价指标	PSO-BP	GA-BP	SOA-BP	SA-SOA-BP
MSE	0.0154	0.0176	0.0072	0.0009
MAPE	0.20%	0.16%	0.16%	0.07%
RMSE	0.1241	0.1326	0.0849	0.0295

其它优化算法得到的衡量指标值都是最小,这从宏观角度表明了基于 SA-SOA-BP 神经网络的网络安全态势预测算法比其它三种算法具有更高的精准性和有效性.

4.3 算法收敛性分析

由于在本实验中是用训练数据预测误差绝对值和作为个体适应度值,个体适应度值越小,说明该个体越优. 因此个体最优适应度的变化情况既可以反应算法的收敛情况,又可以反应训练数据集的训练过程. 图 4 给出了 PSO、GA、SOA 和 SA-SOA 优化 BP 神经网络算法在迭代寻优过程中最优个体适应度值的变化趋势.

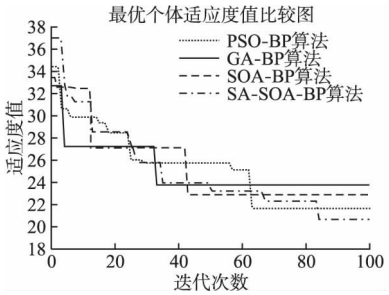

图 4 最优个体适应度值变化对照图

Fig. 4 Comparison chart of optimal individual fitness value changes

从图 4 中可以看出,PSO-BP 算法在开始时得到的最优个体适应度值相对较高,并且自第 25 次和第 62 次迭代时开始长期陷入局部极值,跳出局部极值的时间较长,最优个体适应度值的最小值为 21.727; GA-BP 算法在第 4 次迭代时陷入局部极值,经过 30 次迭代才跳出了局部极值,但是之后又陷入了局部极值,并一直到最后没有再跳出局部极值的陷阱,并且此算法是所有算法中最早陷入局部极值的算法,最优个体适应度值的最小值为 23.837; SOA-BP 算法在前期的优化效果虽然比较好,但是在自第 42 次迭代开始陷入局部极值后,一直陷入局部极值中没有跳出,最优个体适应度值的最小值为 22.957; SA-SOA-BP 算法是最不容易陷入局部最优的算法,多次跳出了局部极值的陷阱,其最优个体适应度值的最小值达到 20.699. 另外在第 84 次迭代时,SA-SOA-BP 算法首先达到了最小的适应度值 20.699,而 SOA-BP 算法在此时并没有达到最小的适应度值,说明了在 100 次迭代内 SA-SOA-BP 算法的收敛速度相对较快达到最小的适应度值. 从图 4 的最优个体适应度值的变化趋势可以看出,采用 SA-SOA 优化 BP 神经网络进行预测的收敛效果比 PSO、GA 和 SOA 算法优化 BP 神经网络进行态势预测的收敛效果更好,其适应度值在趋于平稳时值最小. 综上所述,SA-SOA-BP 算法的收敛性效果相对较好,速度相对较快.

5 总 结

本文将模拟退火算法引入人群搜索算法来优化 BP 神经网络,提出了一种基于 SA-SOA-BP 神经网络的网络安全态势预测方法,解决了 BP 神经网络权值和阈值难以确定和 SOA 算法在搜索后期收敛速度缓慢和易于陷入局部最优等问题,增强了人群搜索算法的全局寻优能力,加快了算法的收敛速度,提高了网络安全态势预测的准确性. 对比实验结果表明,基于 SA-SOA-BP 神经网络的网络安全态势预测算法比 SOA 优化 BP 神经网络、GA 优化 BP 神经网络以及 PSO 算法优化 BP 神经网络的预测结果更为稳定、准确,并具有更好的稳定性和收敛性. 下一步的研究将与其它智能预测算法进行对比,进一步寻找精度和效率更高的网络安全态势预测方法.

References:

[1] Tim Bass. Multisensor data fusion for next generation distributed intrusion detection system [C] //Proceedings of 1999 IRIS National Symposium on Senor and Data Fusion, America: The Johns Hopkins University, 1999: 1-6.

[2] Leau Y B, Manickam S. Network security situation prediction: a review and discussion [C] //Communications in Computer and Information Science, 2015, 516: 424-435.

[3] Husak M, Komarkova J, Bou-Harb E, et al. Survey of attack projection, prediction, and forecasting in cyber security [J]. IEEE Communications Surveys and Tutorials, 2019, 21(1): 640-660.

[4] Fava D S, Byers S R, Yang S J. Projecting cyberattacks through variable-length Markov models [J]. IEEE Transactions on Information Forensics and Security, 2008, 3(3): 359-369.

[5] Fachkha C, Bou-Harb E, Debbabi M. Towards a forecasting model for distributed denial of service activities [C] //Proceedings of the 12th IEEE International Symposium on Network Computing and Applications, 2013: 110-117.

[6] Bhandari P, Singh M. Ontosecure: a semantic web based tool for network security status prediction [C] //Proceedings of the 6th International Advanced Computing Conference, 2016: 551-555.

[7] Hong Xiao-yi. Network security situation prediction based on grey relational analysis and support vector machine [J]. International Journal of Network Security, 2020, 22(1): 177-182.

[8] Jiang Jie, Gao Jia, Chen Tie-ming. Intrusion detection method based on AE-BNDNN model [J]. Journal of Chinese Computer Systems, 2019, 40(8): 1713-1717.

[9] Liu Yu-ling, Feng Deng-guo, Lian Yi-feng, et al. A network security situation prediction method based on time and space dimensional analysis [J]. Journal of Computer Research and Development, 2014, 51(8): 1681-1694.

[10] Huang Tong-qing, Zhuang Yi. A real-time network security situation prediction method [J]. Journal of Chinese Computer Systems, 2014, 35(2): 303-306.

[11] Xu Ya-bin, Jia Shan-shan. Research on security situation awareness of software-defined networks [J]. Journal of Chinese Computer Systems, 2019, 40(8): 1682-1688.

[12] Chen Wei-peng, Zang Zhi-gang, Guo Jie, et al. The security assess-

ment of network space situational awareness system based on improved BP neural network [J]. Computer Science,2018,45(S2): 335-337 +341.

[13] Hu Jing-jing,Ma Dong-yan,Liu Chen,et al. Network security situation prediction based on MR-SVM [J]. Institute of Electrical and Electronics Engineers Inc.,2019,7: 130937-130945.

[14] Han Xiao-lu,Liu Yun,Zhang Zhen-jiang. Network security situation prediction method based on IFS-NARX model [J]. Journal of Jilin University(Engineering and Technology Edition),2019,49(2): 592-598.

[15] Shen Liang,Wen Zhi-cheng. Network security situation prediction in the cloud environment based on grey neural network [J]. Journal of Computational Methods in Sciences and Engineering,2019,19(1): 153-167.

[16] Rumelhart D E,Hinton G E,Williams R J. Learning representations by back-propagating errors [J]. Nature,1986,323(6088): 533-536.

[17] Li Wei,Zhang Chang-sheng. Static friction parameters identification of DC servo system using cuckoo search algorithm based on simulated annealing [J]. Journal of Chongqing University of Posts and Telecommunications(Natural Science Edition),2019,31(6): 892-897.

[18] Li Cao,Liu Xiao-yu. An improved PSO-BP neural network and its application to earthquake prediction [C]//Proceedings of the 28th Chinese Control and Decision Conference(CCDC),2016: 3434-3438.

[19] He Xiao-chuan,Feng Jun-jun,Jia Ru-chun. GABP neural network algorithm applied in evaluation of computer network security [J]. International Journal of Security and its Applications,2016,10(12): 377-388.

[20] Dai C,Zhu Y,Chen W. Seeker optimization algorithm [C]//International Conference on Computational and Information Science,Springer,Berlin,Heidelberg,2006: 167-176.

[21] Metropolis N,Rosenbluth A W,Rosenbluth M N,et al. Equation of state calculations by fast computing machines [J]. Journal of Chemical Physics,1953,21(6): 1087-1092.

[22] Kirkpatrick S,Vecchi M P. Optimization by simulated annealing [M]. Spin Glass Theory and Beyond: an Introduction to the Replica Method and its Applications,1987.

[23] Aylaj B,Belkasmi M,Zouaki H,et al. Degeneration simulated annealing algorithm for combinatorial optimization problems [C]//Proceedings of the 15th International Conference on Intelligent Systems Design and Applications Marrakech,2015: 557-562.

附中文参考文献:

[8] 江　颉,高　甲,陈铁明. 基于 AE-BNDNN 模型的入侵检测方法 [J]. 小型微型计算机系统,2019,40(8): 1713-1717.

[9] 刘玉岭,冯登国,连一峰,等. 基于时空维度分析的网络安全态势预测方法 [J]. 计算机研究与发展,2014,51(8): 1681-1694.

[10] 黄同庆,庄　毅. 一种实时网络安全态势预测方法 [J]. 小型微型计算机系统,2014,35(2): 303-306.

[11] 徐雅斌,贾珊珊. 软件定义网络的安全态势感知研究 [J]. 小型微型计算机系统,2019,40(8): 1682-1688.

[12] 陈维鹏,敖志刚,郭　杰,等. 基于改进的 BP 神经网络的网络空间态势感知系统安全评估 [J]. 计算机科学,2018,45(S2): 335-337 +341.

[17] 李　伟,张长胜. 基于模拟退火布谷鸟算法的直流伺服系统静态摩擦参数辨识 [J]. 重庆邮电大学学报(自然科学版),2019,31(6): 892-897.