

引用格式:张然,刘敏,张启坤,等. 基于 SOA_BP 神经网络的网络安全态势预测算法研究[J]. 微电子学与计算机,2020,37(6):62-65. [ZHANG R,LIU M,ZHANG Q K,et al. Research on network security situation prediction algorithm based on SOA_BP neural network[J]. Microelectronics & Computer,2020,37(6):62-65.]

基于 SOA_BP 神经网络的网络安全态势预测算法研究

张 然¹,刘 敏¹,张启坤¹,甘 勇²

(1 郑州轻工业大学计算机与通信工程学院,河南 郑州 450002;

2 郑州工程技术学院 信息工程学院,河南 郑州 450044)

摘 要:当前网络安全形势日益严峻,为了提高网络安全态势预测的准确性,提出了一种基于人群搜索算法优化 BP 神经网络的网络安全态势预测方法.本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,最后与其它两种优化算法得到的预测值进行对比.实验表明,该算法用于网络安全态势预测精准度更高,误差更小,并具有更好的稳定性.

关键词:BP 神经网络;人群搜索算法;网络安全;态势预测

中图分类号:TP393.08

文献标识码:A

文章编号:1000-7180(2020)06-0062-04

Research on network security situation prediction algorithm based on SOA_BP neural network

ZHANG Ran¹,LIU Min¹,ZHANG Qi-kun¹,GAN Yong²

(1 School of Computer and Communication Engineering,Zhengzhou University of Light Industry,

Zhengzhou 450002, China; 2 School of Information Engineering, Zhengzhou Institute of

Technology, Zhengzhou 450044,China)

Abstract: The current cybersecurity situation is getting worse. In order to improve the accuracy of network security situation prediction, a network security situation prediction method based on SOA_BP neural network is proposed. The algorithm uses the four behavioral characteristics of SOA algorithm: the self-interest, altruism, pre-action and uncertainty reasoning to determine the search strategy, find the best fitness individual, obtain the optimal weight and threshold, and then assign them to the random initial weights and thresholds. After training the neural network, the predicted values are obtained. Finally, it is compared with the predicted values obtained by othertwo optimization algorithms. The experimental results show that this prediction algorithm has higher accuracy, smaller error and better stability

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

1 引言

近些年国内在网络安全态势预测方面也做了大

量的研究,提出了大量的网络安全态势预测模型和方法^[1].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.陆

收稿日期:2019-10-14;修回日期:2019-11-22

基金项目:国家自然科学基金项目(61772477);河南省重点科技攻关项目(142102210081);河南省高等学校重点科研项目计划支持(13B520322);河南省产学研合作项目(132107000066)

冬在其硕士论文[2]中对一种基于遗传算法的 Elman 预测模型进行了研究. 陈维鹏等人[3]提出了一种使用模拟退火算法改进 BP 神经网络的评估模型. 张任川[4]等人,利用卷积神经网络特有的权值共享的特性,并结合深度可分离和可分解的卷积技术,提出了一种基于改进卷积神经网络的网络安全态势预测模型. 钱兆楼[5]等人提出一种基于改进的粒子群算法优化 BP 神经网络的预测方法. 赵伟[6]提出一种将人群搜索算法与支持向量机相结合的网络入侵检测算法. 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.

针对目前安全态势预测精准性不高的问题,本文在已有的网络安全态势预测方法和模型的基础上,将人群搜索算法(SOA)与 BP 神经网络相结合[7],对基于 SOA_BP 神经网络的网络安全态势预测方法进行研究和仿真实验.

2 基于 SOA 算法优化 BP 神经网络的安全态势预测

2.1 人群搜索算法

人群搜索算法(Seeker Optimization Algorithm, SOA)是一种较新的启发式随机搜索算法[8]. 它是对人的随机搜索行为进行分析,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[9]. SOA 算法的优点是概念明确、清晰、易于理解、收敛速度快、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,先将个体最佳适应度值按降序的方法排列,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:

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

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

式中, u_i 表示第 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; $sizepop = 30$ 为种群规模; $U_{\max} = 0.95$, $U_{\min} = 0.0111$ 分别表示最大和最小函数隶属度; u_{ij} 表示为在 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(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:

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

式中, a_{ij} 表示为第 i 个个体在 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.

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

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

式中, $zbest$ 表示为全局最佳; $rand(1, 10)$ 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.

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

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

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

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 个个体当前最佳位置; $\vec{z}_{i,best}$ 表示为全局最佳位置; $F_{i,best}$ 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.

本实验采取的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,计算公式如下:

$$\vec{d}_i(t) = \text{sign}(W \vec{d}_{i,pro} + \varphi_1 \vec{d}_{i,ego} + \varphi_2 \vec{d}_{i,alt}) \quad (9)$$

$$W = W_{\max} - t * \frac{W_{\max} - W_{\min}}{\maxgen} \quad (10)$$

式中, W 为惯性权值; φ_1 和 φ_2 为常数,它们的取值范围为 $[0, 1]$ 内均匀分布的常数. t 为当前迭代次数,

取值范围为 $[2, \max_{gen}]$ 之间的整数, $W_{\max} = 0.9W_{\max} = 0.9$ 为权重最大值, $W_{\min} = 0.1W_{\min} = 0.1$ 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:

$$\Delta x_{ij}(t+1) = \Delta x_{ij}(t) + a_{ij}(t) * d_{ij}(t) \quad (11)$$

2.2 SOA 算法优化 BP 神经网络的预测算法

SOA 算法优化 BP 神经网络(简称 SOA_BP 神经网络)的安全态势预测主要步骤如下:

(1)初始化. 初始化种群规模、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,初始化种群个体,计算出每一个个体的适应度.

(2)寻找具有最好适应度的个体. 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,本步骤最重要的是适应度函数的设计,在此实验里是以 BP 神经网络训练集得到的误差作为适应度值.

(3)迭代寻优. 首先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} . 然后进入迭代循环,循环的终止条件是最大迭代次数和种群规模双层循环. 确定搜索方向,根据公式(6)~(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)~(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(10)更新位置,更新个体最优和最佳适应度值,另外还要更新全局最优和适应度值. 迭代循环直到满足终止条件,输出全局最佳适应度值.

(4)得到最佳网络权值和阈值. 将最优权值和阈值赋给随机初始值用于网络安全训练.

(5)BP 神经网络训练. 初始化网络参数,创建神经网络,调用 traingdm 算法训练 BP 网络,采用 BP 神经网络进行仿真预测,计算误差,分析结果.

SOA 算法优化 BP 神经网络的安全态势预测算法流程如图 1 所示.

3 基于 SOA_BP 神经网络的网络安全态势预测实验仿真与分析

本文将 SOA 算法优化 BP 神经网络的方法应用于网络安全态势预测来提高预测的准确性. 实验以网络安全信息与动态周报中所发布的数据为实验数据,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 and 新增信息安全漏洞数量为评价指标,能够比较

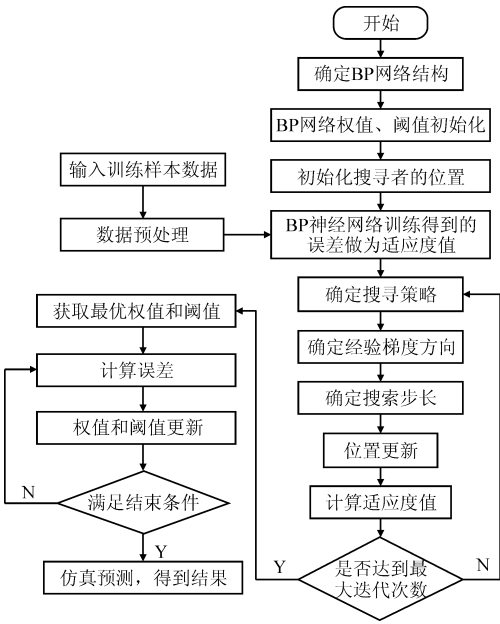


图 1 基于 SOA_BP 的预测算法流程图

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. 在本实验中选取了 2015 年第一期至 2017 年第七期的网络安全数据信息. 为了实验方便,我们将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

优	良	中	差	危
5	4	3	2	1

3.1 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,数量过少将会降低准确度,所以本文选择训练样本数量为 101 个,测试样本数量为 10 个. 为了增加实验的准确性和收敛性,首先对数据进行了预处理和归一化的操作. 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,归一化公式如公式(12),归一化后的结果如图 2 所示.

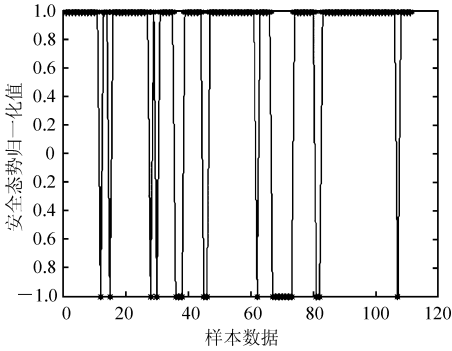


图 2 网络安全标准化值

$$y = 2 \times \frac{x_i - x_{\min}}{x_{\max} - x_{\min}} + (-1) \tag{12}$$

3.2 预测结果分析

(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(13) ~ (15) 确定隐含层节点个数.

$$l < n - 1 \tag{13}$$

$$l < \sqrt{m + n} + a \tag{14}$$

$$l = \log_2 n \tag{15}$$

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 之间的常数. 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.

(2) 验证 SOA 算法的精准性. 采用了均方误差 (mean-square error, MSE) 来衡量真实值与预测值之间的差异. 它是衡量平均误差的一种简单方法, 可以很好的反应数据的变化程度, MSE 越小, 表明此预测的数据越准确, 反之相反. MSE 的计算公式如公式(16).

$$MSE = \frac{1}{n} \sum_{i=1}^n (x_i - x_i^*)^2 \tag{16}$$

式中, x_i 和 x_i^* 分别表示真实值和预测值.

(3) 与其它优化算法的预测结果比对. 下面给出 PSO 优化、GA 优化和 SOA 优化 BP 神经网络的算法实验结果对比图, 如图 3 所示. 从图中可以清楚的

看到各个优化算法真实值与预测值之间的差异.

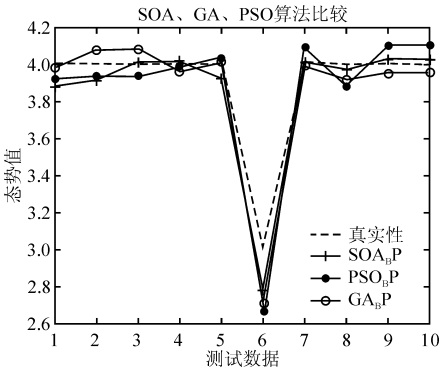


图 3 实验结果对比图

从图 3 可以看出, SOA 优化 BP 神经网络的预测值与真实值之间的误差相对 PSO 优化和 GA 优化 BP 神经网络要小, 并且 SOA 算法优化得到的预测值开始几个与真实值的误差比较大, 但是后面几个逐渐趋于稳定, 接近真实值, 并且曲线波动比较小. 这说明总体上 SOA 优化 BP 神经网络的算法相比其他两种算法更稳定、更准确.

表 2 给出了三个不同优化算法得到的十个测试值以及与真值之间的绝对误差. 表 2 从微观上说明了 SOA_BP 算法的准确性更高.

表 3 分别计算了三个算法的均方误差. 从表 3 可以宏观地看出 SOA_BP 优化算法得到的预测值与真实值的均方误差最小. 这也表明了 SOA 优化算法具有更高的准确性和一定的优越性.

表 2 预测数据分析表

真实值	4	4	4	4	4	3	4	4	4	4
PSO 预测值	3.916	3.934	3.935	3.981	4.030	2.661	4.092	3.881	4.101	4.100
PSO 绝对误差	-0.085	-0.066	-0.065	-0.019	0.030	-0.339	0.092	-0.119	0.101	0.100
GA 预测值	3.980	4.071	4.078	3.959	4.009	2.711	3.991	3.915	3.953	3.953
GA 绝对误差	-0.020	0.071	0.078	-0.041	0.009	-0.289	-0.009	-0.085	-0.047	-0.047
SOA 预测值	3.881	3.912	4.009	4.012	3.924	2.762	4.017	3.968	4.026	4.021
SOA 绝对误差	-0.119	-0.088	0.009	0.012	-0.076	-0.238	0.017	-0.032	0.026	0.021

表 3 精准度对照表

评估指标	PSO_BP	GA_BP	SOA_BP
均方误差	0.017 5	0.010 9	0.008 71

4 结束语

本文针对目前网络安全态势预测的精准性不高的问题, 在利用 SOA 对 BP 神经网络的权值和阈值进行优化的基础上, 提出了一种基于 SOA_BP 神经网络的网络安全态势预测方法, 并将其与 PSO_BP 优化算法和 GA_BP 优化算法进行对比. 实验表明, 此算法能更精准地预测未来的网络安全态势. 但是

SOA 存在搜索后期易陷入局部最优即“早熟”现象, 接下来的工作将对其进一步改进以解决此类缺陷, 进一步提高其全局搜索能力和预测的精准性.

参考文献:

[1] TIM BASS. Multisensordata fusion for next generation distributed intrusion detection system[C]. // In Proceedings of 1999 IRIS National Symposium on Sensor and Data Fusion, America: The Johns Hopkins University, 1999:1-6.

- Proximity-1. Simulating the CCSDS Proximity-1 protocol for ExoMars [C]//Jason T, Data systems; DA-SIA 2009 Data Systems in Aerospace, Istanbul: Data systems publication, 2009: 213-245.
- [2] WANG J, WANG C H, WU H, et al. Parameter Optimization of ARQ System based on Proximity-1 Protocol [C]//WANG Z T, The research of management, computer and medicine: Proceedings of the 2016 7th international conference on education, management, computer and medicine, Shenyang: Springer, 2016: 24-27.
- [3] CCSDS, 211. 0-B-5, Proximity-1 space link protocol - Data link layer [S], Washington DC: CCSDS Secretariat, 2013.
- [4] CCSDS, 211. 0-B-5, Proximity-1 space link protocol - Code and Synchronization sub-layer [S], Washington DC: CCSDS Secretariat, 2013.
- [5] CAPUANO V, BOTTERON C, TIAN J, et al. GNSS to Reach the Moon [C]//Jake R. , IAC: 65th International Astronautical Congress, Toronto: IAF publication, 2014:55-61.
- [6] TIAN J, WANG Y, WANG W, et al. Cross-Band Aided Acquisition on HEO Orbit [C]// Jake R, IAC: 65th International Astronautical Congress, Toronto: IAF publication, 2014:121-130.
- [7] TIAN J, WANG W, WANG D, et al. GPS CA code cross correlation mitigation technique for HEO orbit acquisition [C]// Jake R, IAC: 66th International Astronautical Congress, Jerusalem: IAF publication, 2015: 54-62.
- [8] TIAN J, WANG W, WANG D, et al. Whole MTO orbit acquisition method based on FPGA resource restricted platform [C]// Jake R, IAC: 66th International Astronautical Congress, Jerusalem: IAF publication, 2015: 33-41.
- [9] CCSDS, 211. 0-B-5, CCSDS Proximity-1 space link protocol - Physical layer [S], Washington DC: CCSDS Secretariat, 2013.

作者简介:

田 嘉 男,(1987-),硕士,高级工程师.研究方向为空间通信与导航技术.E-mail:jia_epfl_tian@163.com

王 伟 男,(1970-),硕士,研究员.研究方向为空间通信与导航技术.

史平彦 男,(1964-),硕士,研究员.研究方向为导航与遥感技术.

(上接第 65 页)

- [2] 陆冬.基于 GA-Elman 神经网络的网络态势预测模型研究[D].杭州:浙江大学,2017.
- [3] 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.
- [4] 张任川,张玉臣,刘璟,等.应用改进卷积神经网络的网络安全态势预测方法[J].计算机工程与应用,2019,55(6):86-93.
- [5] 钱兆楼.一种基于改进粒子群算法优化 BP 神经网络的预测方法[J].电子测试,2015(20):39-40.
- [6] 赵伟.基于 SOA-SVM 的网络入侵检测方法研究[J].自动化与仪器仪表,2015(1):39-42.
- [7] RUMELHART D E, HINTON G E, WILLIAMS R J. Learning representations by back-propagating errors [J]. Nature,1986, 323(3):533-536.
- [8] 张连强,王东风.基于改进人群搜索算法的 PID 参数优化[J].计算机工程与设计,2016, 37(12): 3389-3393.
- [9] 高学金,付龙晓,武翠霞,等.基于 ISOA 的 LS-SVM 地铁站空调系统能耗预测模型[J].计算机与现代化,2018(10):36-43.

作者简介:

张 然 女,(1973-),博士,副教授.研究方向为网络信息安全.E-mail:ranranzh@sina.com

刘 敏 女,(1995-),硕士研究生.研究方向为网络信息安全.

张启坤 男,(1980-),博士,副教授.研究方向为信息安全.

甘 勇 男,(1965-),博士,教授.研究方向为网络信息安全.