基于IHS_RELM的网络安全态势预测方法

陈 虹[1] 王 飞[1] 肖振久[1,2]
(辽宁工程技术大学软件学院 葫芦岛 125105)[1] (中国传媒大学计算机学院 北京 100024)[2]

摘 要 针对网络安全态势感知中的态势预测问题,提出一种基于IHS_RELM的网络安全态势预测方法。对和声搜索算法的原理进行了研究,在此基础上提出一种改进的和声搜索算法。将正则极速学习机(RELM)嵌入到改进的和声搜索算法(IHS)的目标函数计算过程中,利用IHS算法的全局搜索能力来优化选取RELM的输入权值和隐含层阈值,在一定程度上提升了RLLM的学习能力和泛化能力。仿真实验表明,与已有的其他预测方法相比,该方法具有更好的预测效果。

关键词 和声搜索算法,正则极速学习机,网络安全态势预测,参数优化
中图法分类号 TP393.08,TP18 **文献标识码** A

Method of Network Security Situation Prediction Based on IHS_RELM

CHEN Hong[1] WANG Fei[1] XIAO Zhen-jiu[1,2]
(School of Software, Liaoning Technical University, Huludao 125105, China)[1]
(School of Computer, Communication University of China, Beijing 100024, China)[2]

Abstract To address the situation prediction problem in the network security situation awareness, this paper presented a prediction method of network security situation based on the algorithm of HIS_RELM. We proposed an improved harmony search(IHS) algorithm after studying the principle of the harmony search(HS) algorithm. This method embeds the regularized extreme learning machine(RELM) in the process of the objective function calculation of the improved harmony search algorithm, and takes advantage of the global searching ability of the IHS algorithm to optimize the input weights and hidden layer threshold of the RELM. To some extent, this enhances the learning ability and generalization ability of the RELM. Simulation experiments show that this method has better prediction affection in comparison with other existing prediction methods.

Keywords Harmony search algorithm(HS), Regularized extreme learning machine(RELM), Network security situation prediction, Parameters optimization

1 引言

随着互联网络的迅速发展,网络安全问题变得越来越严重,网络安全主动防御策略也成为当前网络安全领域的研究热点之一。文献[1]中Songmei Zhang等提出一种基于信息融合的网络安全态势分析框架,力图重现网络遭受攻击的过程;文献[2]中韩敏娜等提出一种基于集对分析的网络安全态势评估方法,利用集对分析理论评估网络安全态势值。但文献[1,2]都没能实现网络安全态势预测。

文献[3]中孟锦等提出基于HHGA-RBF神经网络的网络安全态势预测模型;文献[4]中尤马彦等提出基于Elman神经网络的网络安全态势预测方法;文献[5]中王晋东等提出使用马尔可夫链结合灰色理论构造预测模型。但这些方法都有不足之处:文献[3,4]采用的神经网络方法具有结构难以确定和易陷入局部最优的缺点;文献[5]中提出的数学模型在实际应用中难以建立,而且需要大量的复杂数学推理过程。

正则极速学习机[6](RELM)是经典极速学习机(ELM)的改进和发展,它借鉴统计学习理论中结构风险最小化原理,引入参数γ来调节经验风险与结构风险,提高了ELM的泛化能力。但RELM的输入权值和隐含层阈值是随机生成的,这可能导致所得解在实际应用中并不是参数最优的。和声算法[7]是新近提出的一种全局优化算法,本文对其进行了改进,并将改进后的和声搜索算法用于优化选取RELM的输入权值和隐含层阈值,提出一种IHS_RELM算法。最后尝试将该算法应用到网络安全态势预测中,同时与已有模型进行对比。仿真实验表明,该方法更具优越性。

2 RELM算法

给定一时间序列数据集:$G=\{(x_1,t_1),\cdots,(x_N,t_N)\}$,其中$x_i=[x_{i1},\cdots,x_{in}]\in R^n$,$t_i=[t_{i1},\cdots,t_{in}]\in R^m$,$i=1,2,\cdots,N$,一个隐含层节点数为$L$、激励函数为$g(x)$的RELM回归模型为:

到稿日期:2013-01-21 返修日期:2013-06-05 本文受国家自然科学基金(61103199)资助。
陈 虹(1967-),女,副教授,主要研究领域为网络安全;王 飞(1988-),男,硕士生,主要研究领域为网络安全,E-mail:china_wangfei@163.com;肖振久(1968-),男,副教授,主要研究领域为网络与信息安全、数字版权管理。

$$\sum_{i=1}^{L}\beta_i g(\alpha_i \cdot x_j + b_i) = t_j, j=1,2,\cdots,N \quad (1)$$

式中，$\alpha_i=[\alpha_{i1},\alpha_{i2},\cdots,\alpha_{in}]^T$ 是连接第 i 个隐含层节点的输入权值，b_i 是第 i 个隐含层节点的阈值，$\beta_i=[\beta_{i1},\beta_{i2},\cdots,\beta_{im}]^T$ 是连接第 i 个隐含层节点的输出权值，$\alpha_i \cdot x_j$ 表示 α_i 和 x_j 的内积，激励函数 $g(x)$ 可以是“Sine”、“Sigmoid”或“RBF”等。

式(1)中包含 N 个方程，可写成如下的矩阵形式：

$$H\beta = T \quad (2)$$

式中，H 表示隐含层输入矩阵，具体形式为：

$$H(\alpha_1,\alpha_2,\cdots,\alpha_L,b_1,b_2,\cdots,b_L,x_1,x_2,\cdots,x_N)=$$

$$\begin{bmatrix} g(\alpha_1 \cdot x_1+b_1) & g(\alpha_2 \cdot x_1+b_2) & \cdots & g(\alpha_L \cdot x_1+b_L) \\ g(\alpha_1 \cdot x_2+b_1) & g(\alpha_2 \cdot x_2+b_2) & \cdots & g(\alpha_L \cdot x_2+b_L) \\ \vdots & \vdots & \vdots & \vdots \\ g(\alpha_1 \cdot x_N+b_1) & g(\alpha_2 \cdot x_N+b_2) & \cdots & g(\alpha_L \cdot x_N+b_L) \end{bmatrix}_{N\times L}$$

$$\beta=\begin{bmatrix}\beta_1^T\\ \beta_2^T\\ \vdots\\ \beta_L^T\end{bmatrix}_{L\times m}, T=\begin{bmatrix}t_1^T\\ t_2^T\\ \vdots\\ t_N^T\end{bmatrix}_{N\times m}$$

与 ELM 回归模型不同，RELM 回归预测即是求解如下的优化问题：

$$\arg\min\left(\frac{1}{2}\|\beta\|^2+\frac{\gamma}{2}\|\varepsilon\|^2\right) \quad (3)$$

$$\text{s.t.}\ \sum_{i=1}^{L}\beta_i g(\alpha_i \cdot x_j+b_i)-t_j=\varepsilon_j, j=1,2,\cdots,N$$

式中，误差平方和 $\|\varepsilon\|^2$ 代表经验风险，$\|\beta\|^2$ 代表结构风险，γ 是调节 2 种风险的参数。

为求解上述优化问题，构造如下拉格朗日函数：

$$L(\beta,\varepsilon,\omega)=\frac{1}{2}\|\beta\|^2+\frac{\gamma}{2}\|\varepsilon\|^2-\omega(H\beta-T-\varepsilon) \quad (4)$$

式中，$\omega=[\omega_1,\omega_2,\cdots,\omega_N]$ 表示拉格朗日乘子。

根据拉格朗日条件，对拉格朗日函数求偏导数并令偏导数为零，可以求得：

$$\beta=\left(H^TH+\frac{I}{\gamma}\right)^{-1}H^TT \quad (5)$$

式中，I 为单位矩阵。最终可得 RELM 时间序列预测模型如下：

$$t=\sum_{i=1}^{L}\beta_i g(\alpha_i \cdot x+b_i) \quad (6)$$

式中，x 为模型输入向量，t 为模型输出向量。

3 改进的和声搜索算法

3.1 和声搜索算法(HS)

HS算法首先随机产生 HMS(Harmony memory size, HMS)个初始解(和声)放入和声记忆库(Harmony memory, HM)中，根据相应规则产生新解，然后判断新解是否优于 HM 内的最差解，若是则替换之，否则保持当前 HM 不变。上述过程不断重复，直至满足终止条件为止。

HS算法过程如下：

Step1 HM 的初始化包括以下两部分：

1)初始化算法参数。包括和声记忆库大小 HMS、和声记忆库考虑概率 HMCR、和声微调概率 PAR、微调幅度 BW 和算法迭代次数 NI。

2)初始化和声记忆库。初始和声矢量 $X_1,\cdots,X_{HMS}$ 在定义域内按照 $X_{i,j}=LB_j+r\times(UB_j-LB_j)$ 均匀地产生，其中 $X_{i,j}$ 为 X_i 的第 j 维决定变量，r 为[0,1]之间的均匀随机数，UB_j 和 LB_j 分别为决策变量的上界和下界。

Step2 基于考虑概率、微调概率和随机选择 3 个规则产生新的和声矢量。过程描述如下：

```
for(j=1;j<=N;j++)
{ if(r1<HMCR)//r1=rand(0,1);
  { Xnew,j=Xi,j;//i 为[1,HMCR]之间的随机数
    if(r2<PAR) //r2=rand(0,1);
      Xnew,j=Xnew,j± r2×BW;
  }else
  {Xnew,j= LBj+r3×(UBj-LBj);//r3=rand(0,1);
  }
}
```

Step3 更新和声记忆库。令 $f(x)$ 表示目标函数，$f(x)$ 值越小表示性能越好。如果 $f(X_{new})<f(X_{worst})$，则 $X_{worsi}=X_{new}$，并将 HM 按照 $f(x)$ 的优劣重新排序；否则，保持 HM 不变。

Step4 判断是否满足终止条件，若是则算法结束，否则转到 Step2 继续执行。

3.2 改进和声搜索算法(IHS)

通过对 3.1 节分析可知，和声记忆库考虑概率 HMCR、和声微调概率 PAR 和微调幅度 BW 是和声搜索算法的 3 个关键控制参数。根据文献[8]，HMCR 值大有利于算法局部收缩，值小有利于群体多样性，本文取 HMCR 值为 0.95；PAR 值大有利于算法在和声记忆库中调整搜索区域，值小有利于算法增强局部搜索能力；BW 值大有利于算法跳出局部最优，值小有利于算法在局部区域精细搜索。在基本 HS 算法中，PAR 和 BW 的值在整个迭代过程中是固定不变的，这严重影响了算法的性能。为了在整个解空间进行有效搜索，并尽可能将搜索重点集中于性能较高的区域，从而提高算法效率，本文采用动态变化的 PAR 和 BW，即 PAR 值由小到大变化，BW 值由大到小变化。

3.2.1 设置 PAR 变化方式

在 HS 算法搜索初期，采用较小的 PAR 值有利于算法快速搜寻较好区域；在 HS 算法搜索后期，采用较大的 PAR 值有利于算法跳出局部极值。因此，首先确定 PAR 值的变化范围，然后采用从小到大的变化方式。PAR 值按式(7)动态变化，如图 1 所示。

$$PAR(iter)=PAR_{\min}+(PAR_{\max}-PAR_{\min})\times \sin\left(\frac{\pi}{2}\cdot\frac{iter}{NI}\right) \quad (7)$$

式中，$PAR_{\max}$ 和 $PAR_{\min}$ 分别为 PAR 的上界和下界，iter 和 NI 分别为当前和最大迭代次数。

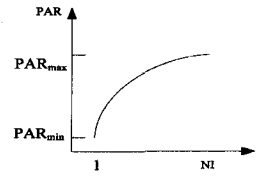

图 1 PAR 值的正弦变化

起始时 PAR 为最小值，随着迭代次数的增加，PAR 值按正弦曲线逐渐增大，当迭代次数 iter 趋近于 NI 时，PAR 趋近

于最大值。

3.2.2 设置 BW 变化方式

在 HS 算法搜索初期，采用较大的 BW 值有利于算法在大范围内探测；在 HS 算法搜索后期，采用较小的 BW 值有利于算法在小范围内精细搜索。因此，首先确定 BW 的变化范围，然后采用从大到小的变化方式。BW 值按式(8)动态变化，如图 2 所示。

$$BW(iter)=BW_{\min}+(BW_{\max}-BW_{\min})\times \cot\left(\frac{\pi}{4}+\frac{\pi}{4}\cdot\frac{iter}{NI}\right) \tag{8}$$

式中，$BW_{\max}$ 和 $BW_{\min}$ 分别为 BW 的上界和下界，iter 和 NI 分别为当前和最大迭代次数。

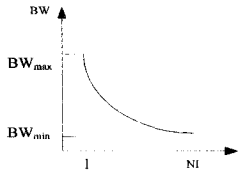

图 2 BW 值的线性变化

起始时 BW 为最大值，随着迭代次数的增加，BW 值按余切方式逐渐减小，当迭代次数 iter 趋近于 NI 时，BW 趋近于最小值。

4 IHS_RELM 预测方法

由文献[6]可知，RELM 的输入权值和隐含层阈值在训练过程中随机产生，且在整个过程之中无需调整，这将影响隐含层输出矩阵中每个元素的取值，进而影响到由解析法确定的输出权值的取值。

对于如何获取最佳的 RELM 输入权值和隐含层阈值，目前还没有统一的方法。文献[9]提出一种基于差分进化优化 ELM 的模拟电路故障诊断的方法，这种方法需要人为确定变异因子、交叉因子和选择因子 3 个参数，并且要求激活函数是无限可微的。为了使 RELM 获得更好的泛化性能和预测精度，本文使用改进的 HS 算法对 RELM 算法的输入权值和隐含层阈值进行优化选择，提出一种 IHS_RELM 算法。在 IHS_RELM 算法中，每个和声代表 RELM 的一组输入权值和阈值，和声所对应的目标函数值反映了该组参数下的算法性能，本文选取均方根误差(RMSE)作为目标函数，其形式如下：

$$f_{RMSE}=\sqrt{\frac{1}{N}\sum_{i=1}^{N}(t_i-t_i')^2} \tag{9}$$

式中，N 是训练样本的个数，t_i 是实际值，t_i' 是预测值，f_{RMSE} 是相应的目标函数值。

IHS_RELM 算法过程如下：

Step1 初始化算法参数，随机生成并初始化和声库，其中每个和声代表一组输入权值和阈值($\alpha_1,\alpha_2,\cdots,\alpha_L,b_1,b_2,\cdots,b_L$)；

Step2 基于考虑概率、微调概率和随机选择 3 个规则产生新的和声，其中 PAR 和 BW 的值按式(7)和式(8)进行动态变化；

Step3 根据式(6)和式(9)更新和声记忆库，若新解优于 HM 内的最差解，则替换之，并将 HM 内各个解向量重新排序，否则保持当前 HM 不变。

Step4 判断是否满足终止条件或达到预定迭代次数 NI，若是则转到 Step5，否则转到 Step2 继续执行。

Step5 输出和声库 HM 的最优解向量，按照式(6)构造最优的回归预测模型。

5 仿真实验

5.1 实验数据及其相关处理

本文实验采用 Honeynet 组织收集的黑客攻击数据[10]，采用文献[11]提出的网络安全态势评估方法来计算 2000 年 7 月 5 日到 2000 年 12 月 3 日的网络安全态势值。这期间的记录数据相对完整，共得到 126 个网络安全态势值。为了避免原始数据跨度大对预测模型训练造成不良影响，将所获得的样本数据集归一化到区间(0,1)。归一化公式如下所示：

$$x_i'=\frac{x_i-x_{\min}}{x_{\max}-x_{\min}},i=1,2,\cdots,n \tag{10}$$

式中，x_i 与 x_i' 分别为归一化前后的网络安全态势值；$x_{\max}$ 与 $x_{\min}$ 分别为归一化前所有网络安全态势值中的最大值和最小值；n 为网络安全态势值个数。归一化后的网络安全态势值如图 3 所示。

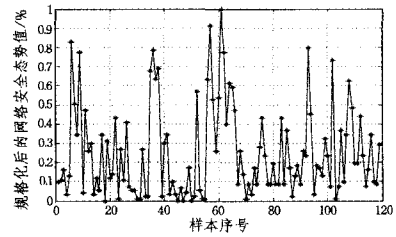

图 3 规格化后的网络安全态势值

5.2 实验预测分析

由于网络安全态势值都是一维的时间序列值，需要重构这些值才能得到符合条件的样本集。重构就是确定输入维数和输出维数的过程，本文设定输入维数为 7，输出维数为 1。对归一化后的样本数据集进行重构，可以构造 119 个样本对，选取前 105 个样本对作为训练集，后 14 个样本对作为测试集。

本文所有实验的环境为：Windows xp 操作系统，matlab2010a 平台，Core CPU 主频为 2000Hz，RAM 为 2048MB。利用第 4 节的方法建立网络安全态势预测模型，对测试集数据进行实验。为了验证本文算法有更好的预测精度，采用已有的网络安全态势预测方法(文献[3]的 HHGA-RBFNN 预测方法、文献[4]的 Elman 预测方法)进行相同的实验，得到如图 4 所示的对比结果。

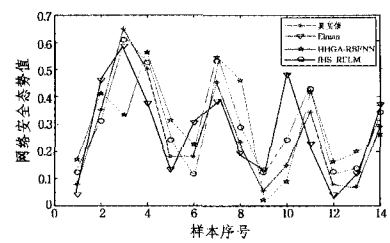

图 4 Elman、HHGA-RBFNN、IHS_RELM 态势预测值

从图 4 中可以看出 IHS_RELM 模型的预测结果要优于其他两种方法，下面采用定量分析方法作进一步的对比。

本文采用均方根误差(RMSE)和平均相对误差(MAPE)两项性能指标来评判预测模型的性能。RMSE 和 MAPE 值越小,对应的模型预测性能越好。这 3 种模型的性能对比如表 1 所列。

表 1　3 种模型的性能对比

预测模型	RMSE	MAPE
IHS_RELM	0.024193	0.52617
Elman	0.049287	0.58374
HHGA_RBFNN	0.050416	0.60159

从表 1 中可以看出,IHS_RELM 预测模型的 RMSE 和 MAPE 值均小于其他两种模型的 RMSE 和 MAPE 值,表明 IHS_RELM 模型的预测性能优于其他两种模型。

结束语　对网络安全态势进行预测是主动防御黑客攻击的一种有效手段,有助于网络管理人员把握未来网络安全态势的发展趋势,从而提前采取相应的网络安全措施。本文提出一种基于 IHS_RELM 的网络安全态势预测方法,其将 RELM 嵌入到 IHS 算法的适应度计算过程中,利用 IHS 算法的全局搜索能力来优化选取 RELM 的输入权值和隐含层阈值,在一定程度上提升了 RELM 的学习能力和泛化能力。仿真实验结果表明,该方法对于预测未来的网络安全态势值具有较好的效果。

下一步的研究工作:

1. 如何获取 RELM 算法中最佳的隐含层节点数,期望获取更好的模型结构;

2. 如何将 IHS_RELM 算法与在线学习算法结合起来,期望实现实时网络安全态势预测。

参 考 文 献

[1] Zhang Song-mei, Yao Shan, Ye Xin'en, et al. A Network Security Situation Analysis Framework Based on Information Fusion [C]// Information Technology and Artificial Intelligence Conference(ITAIC). 2011 6th IEEE Joint International. 2011, 1: 326-332

[2] 韩敏娜,刘渊,陈烨. 基于集对分析的网络安全态势评估[J]. 计算机应用研究,2012,29(10):3824-3827

[3] 孟锦,马驰,何加浪,等. 基于 HHGA-RBF 神经网络的网络安全态势预测模型[J]. 计算机科学,2011,38(7):71-75

[4] 尤马彦,凌捷,郝彦军. 基于 Elman 神经网络的网络安全态势预测方法[J]. 计算机科学,2012,39(6):61-76

[5] 王晋东,沈柳青,王坤,等. 网络安全态势预测及其在智能防护中的应用[J]. 计算机应用,2010,30(6):1480-1488

[6] 邓万宇,郑庆华,陈琳,等. 神经网络极速学习方法研究[J]. 计算机学报,2010,33(2):279-287

[7] Geem Z W, Kim J H, Loganathan G V. A new heuristic optimiza-tion algorithm: harmony search[J]. Simulation, 2001, 76(2): 60-68

[8] Omran M G H, Mahdavi M. Global-best Harmony Search[J]. Applied Mathematics and Computation, 2008, 198(2): 643-656

[9] 周江嫚,黄清秀,彭敏放,等. 基于差分进化优化 ELM 的模拟电路故障诊断[Z]. 计算机工程与应用,2012

[10] Honeynet Project. Know Your Enemy: Statistics[EB/OL]. http://old.honeynet.org/papers/stats/, 2001

[11] 陈秀真,郑庆华,管晓宏,等. 层次化网络安全威胁态势量化评估方法[J]. 软件学报,2006,17(4):885-897

(上接第 97 页)

$nl(f)$的下界范围更广了。所以构造 2 中的函数不仅在形式上具有普遍性,性质上也同样具有一般性。

结束语　本文给出了一种具有最优代数免疫度的偶数元布尔函数的构造,并且还给出了一种具有最优代数免疫度的平衡偶数元布尔函数的构造。因为猜想 4 和猜想 3 等价,我们也可以从猜想 3 的角度出发进行较特殊情况的构造,用 xy^u 代替 x^vy^u,经计算证明发现也可得到相同结果,且计算更简便。本文中还存在一些亟待解决的问题,比如当 u,v 为何值时,构造 1 中的函数为 Bent 函数;所构造函数与其它已知函数性质的比较等等,都是我们下一步要研究的重点。

参 考 文 献

[1] Armknecht F. Improving fast algebraic attacks: FSE 2004[C]// LNCS 3017. Springer Verlag, 2004: 65-82

[2] Batten L M. Algebraic attacks over GF(q): Cryptology-INDOCRYPT 2004[C]// LNCS 3348. Springer Verlag, 2004: 84-91

[3] Courtois N, Meier W. Algebraic attacks on stream ciphers with linear feedback: Cryptology-EUROCRYPT 2003[C]// LNCS 2656. Springer Verlag, 2003: 345-359

[4] Courtois N. Fast algebraic attacks on stream ciphers with linear feedback: Advances in Cryptology-CRYPTO 2003[C]// LNCS 2729. Springer Verlag, 2003: 176-194

[5] Meier W, Pasalic E, Carlet C. Algebraic attacks and decomposition of Boolean functions: Cryptology-EUROCRYPT 2004[C]// LNCS 3027. Springer Verlag, 2004: 474-491

[6] 孟强,陈鲁生,符方伟. 一类代数免疫度达到最优的布尔函数的构造[J]. 软件学报,2010: 1758-1767

[7] 涂自然,邓映蒲. 代数免疫度为 1 的布尔函数[J]. 系统科学与数学,2011,31(5): 512-518

[8] 李超,薛朝红,付绍静. 代数免疫度最优的旋转对称布尔函数的构造[J]. 国防科技大学学报,2012,34(2): 34-38

[9] Tu Z, Deng Y. A conjecture about binary strings and its applications on constructing Boolean functions with optimal algebraic immunity: Des[J]. Codes Cryptogr, 2011, 60(1): 1-14

[10] Tang D, Carlet C, Tang X. Highly nonlinear Boolean functions with optimum algebraic immunity and good behavior against fast algebraic attacks[J]. Cryptology ePrint Archive, 2013, 59(1): 653-664

[11] Jin Q, Liu Z, Wu B, et al. A general conjecture similar to T-D conjecture and its applications in constructing Boolean functions with optimal algebraic immunity[C]// Cryptology ePrint Archive 2011. 2011: 515

[12] Lidl R, Niederreiter H. Finite Fields, Encyclopedia of Mathematics and its Applications[M]. 1983

[13] Carlet C, Feng K. An infinite class of balanced functions with optimal algebraic immunity, good immunity to fast algebraic attacks and good nonlinearity: Asiacrypt 2008[C]// LNCS 5350. SpringernVerlag, 2008: 425-440