**Computer Science and Application** 计算机科学与应用, 2018, 8(11), 1650-1656
Published Online November 2018 in Hans. http://www.hanspub.org/journal/csa
https://doi.org/10.12677/csa.2018.811181

Hans汉斯

# Network Intrusion Detection Model Based on Bat Optimization Algorithm

**Qingjie Zhao[1], Longge Wang[1*], Jie Li[1], Junyang Yu[1,2]**

[1]College of Software, Henan University, Kaifeng Henan
[2]State Key Laboratory of Network and Exchange Technology, Beijing University of Posts and Telecommunications, Beijing
Email: su__xiu@163.com, *wlg@henu.edu.cn

Received: Oct. 15th, 2018; accepted: Oct. 26th, 2018; published: Nov. 2nd, 2018

## Abstract

**Network intrusion has the characteristics of sudden and concealment, and the traditional technology is difficult to describe the law of change, which leads to a very low accuracy of intrusion detection. In order to improve the accuracy of intrusion detection and reduce the false detection rate, a bat optimization algorithm based on dynamic adaptive weight and Cauchy mutation was proposed to optimize the neural network intrusion detection model. It is necessary to collect the data of the intrusion network and then import the data into the neural network to learn. The bat optimization algorithm is used to optimize the parameters of the network model. Finally, the KDD CUP 99 dataset is selected to simulate the network intrusion detection. The results show that the proposed model can obtain ideal network intrusion detection rate and false detection rate.**

## Keywords

**Bat Optimization Algorithm, Neural Network, Intrusion Detection Model, Model Parameters**

# 基于蝙蝠优化算法的网络入侵检测模型

**赵青杰[1]，王龙葛[1*]，李　捷[1]，于俊洋[1,2]**

[1]河南大学软件学院，河南　开封
[2]北京邮电大学网络与交换技术国家重点实验室，北京
Email: su__xiu@163.com, *wlg@henu.edu.cn

收稿日期：2018年10月15日；录用日期：2018年10月26日；发布日期：2018年11月2日

## 摘　要

**网络入侵具有突发性和隐蔽性等特点，传统的技术很难描述其变化规律，这导致入侵检测正确率非常的**

*通讯作者。

文章引用: 赵青杰, 王龙葛, 李捷, 于俊洋. 基于蝙蝠优化算法的网络入侵检测模型[J]. 计算机科学与应用, 2018, 8(11): 1650-1656. DOI: 10.12677/csa.2018.811181

**低。为了提高入侵检测正确率，降低误检率，提出了一种基于动态自适应权重和柯西变异的蝙蝠优化算法优化神经网络的入侵检测模型。需要先采集入侵网络的数据进行整理，然后导入到神经网络中学习，采用蝙蝠优化算法优化网络模型的参数。最后选取KDD CUP 99数据集进行网络入侵检测的仿真实验。结果表明，本文模型能够获得理想的网络入侵检测率和误检率。**

## 关键词

**蝙蝠优化算法，神经网络，入侵检测模型，模型参数**



## 1. 引言

21 世纪是网络和信息化的时代，“ABCD”(分别代指人工智能、区块链技术、云端服务和大数据技术)正在成为信息产业进一步发展的方向和当下的科技风口，新技术新发明的出现也使得我们的日常生活发生了翻天覆地的变化。当然这一切的技术似乎都离不开网络的支撑。计算机网络从上个世纪诞生以来一直都在飞速的发展，现在的互联网技术早已进入千家万户，融入了人们日常生活的方方面面，成为人们从事生产生活或者娱乐学习的有力工具，人们对于网络的依赖也越来越深。但随着网络规模的不断扩大，网络入侵事件也日益增多。网络一旦受到恶意入侵和攻击，损失将会是惨重的。入侵检测系统一直以来都是网络安全研究领域的一个热点[1]，它作为网络安全防御的一条重要防线，能够检测出各种入侵行为，有力的保障用户数据安全和网络正常。

针对网络入侵检测的问题，国内外许多专家学者进行过很多研究，同时也提出了很多有效的网络入侵检测模型。要进行入侵检测需要收集各种状态的网络数据，这其中可能有正常的和非正常的，对这些数据进行系统的分析，从而判断网络行为，然后根据不同的网络行为采取相应的安全举措[2]。反向传播(Back Propagation，BP)神经网络是一种非线性的同时具有很强分类能力的神经网络，具有简单、易实现、自组织和适应能力强等优点，成为目前应用最为广泛的网络入侵检测方法[3] [4]。回声状态网络(Echo State Network，ESN)是一种新型的人工神经网络，一经提出便成为学术界研究的热点，并被大量地应用到了各种不同的领域，包括机器人控制、动态模式分类、事件监测等，尤其是在网络入侵的检测问题上，ESN 也取得了较为突出的效果。蝙蝠算法(Bat Algorithm，BA)是一种模拟自然界蝙蝠回声定位现象而开发的一种新型群智能算法，相比较于遗传算法、粒子群算法等传统智能算法优化的入侵检测模型[5] [6]，蝙蝠算法能够发挥更大的作用。蝙蝠算法可以动态控制局部搜索和全局搜索，具有较好的收敛性，且能避免陷入局部最优，其为神经网络的参数优化问题提供了一种新的研究思路和研究方法[7] [8] [9] [10]。

在网络入侵检测的过程中，网络入侵分类器的设计是网络入侵检测系统的关键，当前网络入侵分类器主要是基于 K 最近邻算法、支持向量机、神经网络等进行的设计[11] [12] [13] [14]。其中出回声状态神经网络是一种新型神经网络，它具有简单、易实现、泛化能力强等优点，成为众多网络入侵检测研究中的一个重要的研究方向[15]。基于动态自适应权重和柯西变异的蝙蝠优化算法(Bat Optimization Algorithm Based on Dynamically Adaptive Weight and Cauchy Mutation, BOA)是一种对基本蝙蝠算法改进优化后的群智能搜索算法，该算法在速度公式中加入了动态自适应权重，可以动态地调整自适应权重的大小，加快算法的收敛速度。此外，通过引入柯西逆累积分布函数方法，在每次迭代时，能有效提高蝙蝠算法的全局

搜索能力，避免陷入局部最优。该算法可以为回声状态神经网络的参数优化问题提供一种新的解决思路。

## 2. 基于蝙蝠优化算法的网络入侵检测模型

### 2.1. 回声状态神经网络

ESN 是由输入层、隐藏层(即储备池)、输出层组成的一种新型的非线性递归神经网络[16]。其将隐藏层设计成了一个由很多神经元组成的稀疏网络，通过调整网络内部权值的特性达到记忆数据的功能，其内部的动态储备池包含了大量稀疏连接的神经元，包含了系统的运行状态，并具有短期训练记忆的功能。其状态方程为：

$$x(t+1)=sigmoid\left[\alpha\cdot Win\cdot u(t)+\beta\cdot Wx\cdot x(t)\right] \tag{1}$$

上式中，$sigmoid$ 为激活转换函数；$\alpha$ 为输入项的比例系数；$\beta$ 为内部储备池的谱半径，$Win$ 和 $Wx$ 分别为输入和储备池内部的连接矩阵；$u(t)$，$x(t)$ 则分别表示 $t$ 时刻的输入向量和储备池内部的状态向量。

那么 ESN 的输出方程为

$$y(t)=x(t)\cdot Wout \tag{2}$$

式中，$y(t)$ 为 $t$ 时刻的输出向量，$Wout$ 为输出连接向量。

ESN 训练的过程，就是训练隐藏层到输出层的连接权值($Wout$)的过程。输出权值对 ESN 的性能起着非常关键的作用，常采用最小二乘法进行求解[17]，目标函数最小化的形式为：

$$\min\left\|X\cdot Wout-Y\right\| \tag{3}$$

式中，$Y=\left[y(1),y(2),\cdots,y(i)\right]^{\mathrm{T}}$，$X=\left[x(1),x(2),\cdots,x(i)\right]^{\mathrm{T}}$，$X\in R_i\cdot N$，$N$ 为储备池的节点数；$i$ 为训练样本数。

根据式(3)得到下面解：

$$West=X+Y=\left(X^{\mathrm{T}}X\right)^{-1}X^{\mathrm{T}}Y \tag{4}$$

式中，$West$ 为 $Wout$ 的估计值。

从(1)式可以看出，参数 $\alpha$ 和 $\beta$ 的选取影响回声状态神经网络的性能。本文采用基于动态自适应权重和柯西变异的蝙蝠优化算法(BOA)对参数 $\alpha$ 和 $\beta$ 进行选取，以达到提高网络入侵检测模型检测正确率的目的。

### 2.2. 基于动态自适应权重和柯西变异的蝙蝠优化算法

基本蝙蝠算法的速度更新系数为 1，这样的设置使得速度更新方式过于单一，容易使蝙蝠种群陷入僵化状态，导致蝙蝠不能动态地寻找猎物，减少了蝙蝠种群的多样性。从而，不能使蝙蝠种群在寻优过程中协调好局部搜索和全局探索之间的平衡关系。为了解决该问题，受文献[18]的启发，本文提出了新的动态自适应权重函数，其数学表达式为：

$$w=\cos\left(\frac{\pi\cdot t}{2\cdot T_{\max}}+\pi\right)+1 \tag{5}$$

其中，$t$ 为当前迭代次数，$T_{\max}$ 为最大迭代次数。这样，通过引入余弦函数使蝙蝠速度更新方式更具灵活性，改变了速度更新系数恒定为 1 的僵化状态。这样蝙蝠算法的速度更新系数就由固定的值 1 变成了在[0, 2]之间动态变化取值。最后，加上 1 可以使蝙蝠速度整体提高，又由于余弦函数最大值为 1 的缘故，不致使蝙蝠寻觅猎物的飞行速度过快。从而，使全局搜索和局部搜索得到平衡。据此，速度更新公式变为

$$V_{\mathrm{i}}^{t}=w\cdot V_{i}^{t-1}+\left(x_{i}^{t}-x_{*}\right)\cdot f_{i} \tag{6}$$

柯西分布也叫作柯西－洛伦兹分布，以奥古斯丁－路易－柯西与亨德里克－洛伦兹名字命名，它是一个数学期望不存在的连续型概率分布。当随机变量 X 满足其概率密度函数时，称 X 服从柯西分布。柯西分布具有原点处概率密度大分布紧凑，而两端密度小分布较长的特点。蝙蝠种群经过柯西变异可以在当前变异个体附近生成更大的扰动，使得蝙蝠可以在更广的范围内更新位置进而跳出局部极值。受文献[19]的启发，本文选取了柯西逆累积分布函数对蝙蝠进行变异。利用柯西分布“尾巴长”的特点，使蝙蝠个体朝更广的范围变异，这样能尽量避免蝙蝠算法陷入局部的最优解。柯西逆累积分布函数如公式(7)所示。将蝙蝠算法中位置的迭代更新公式改成公式(8)。

$$F^{-1}\left(p;x_{0},\gamma\right)=x_{0}+\gamma\cdot\tan\left(\pi\cdot\left(p-\frac{1}{2}\right)\right) \tag{7}$$

$$X_{i}^{t}=X_{i}^{t-1}+V_{i}^{t}\cdot\boldsymbol{A}\cdot\tan\left(\pi\cdot\left(r-\frac{1}{2}\right)\right) \tag{8}$$

式中，$F^{-1}$ 是柯西分布的逆累积分布函数，$X_{i}^{t-1}$ 是变异前的第 $i$ 只蝙蝠的位置点，$\gamma=\boldsymbol{A}$，$r\in[0,1]$ 的均匀分布。其中 $\boldsymbol{A}$ 为系数向量，数学表达式如(9)：

$$\boldsymbol{A}=2\cdot\boldsymbol{a}\cdot\boldsymbol{r} \tag{9}$$

式中，$\boldsymbol{A}$ 是从 2 到 0 线性递减的向量，$\boldsymbol{r}\in[0,1]$ 的均匀随机向量。

### 2.3. BOA 算法优化神经网络参数

优化流程如下：

1) 采集网络入侵的数据并进行清洗整理。

2) 初始化蝙蝠群：蝙蝠个体数为 $n$；每只蝙蝠的音量 $A_{i}^{t}$ 和脉冲频率 $r_{i}$；蝙蝠的频率范围 $f$；蝙蝠的位置 $x_{i}$；蝙蝠的速度 $v_{i}$ 以及迭代次数和误差精度。

3) 初始位置为 ESN 的参数；导入训练样本和测试样本。

4) 根据适应度函数，求出蝙蝠个体的适应度函数值，找出最优结果并记录最优值的蝙蝠的位置。

5) 利用式 6)~8)对蝙蝠的搜索速度、位置和脉冲频率进行更新。得到每只蝙蝠的最佳位置和整体的最佳位置。

6) 生成均匀分布随机数 rand，如果 $\mathrm{rand}<A_{i}$ 且 $f\left(x_{i}\right)<f\left(x_{*}\right)$，则接受步骤 2)产生的新解，然后按响度更新公式和脉冲发射率更新公式对 $A_{i}^{t}$ 和 $r_{i}$ 进行更新。

7) 计算所有蝙蝠的适应度值并对其进行排序，得到全局最佳位置的适应度值。若达到网络训练精确度或当前最大迭代次数，则迭代结束转至步骤 8)；否则计算各蝙蝠的个体极值和全局极值的位置，继续更新蝙蝠的速度和位置。

8) 将蝙蝠全局最优位置对应的值作为回升状态神经网络的参数，并利用最优参数建立网络入侵检测模型。

## 3. 仿真实验

### 3.1. 仿真环境

本实验的实验数据来自 KDD CUP 1999 网络入侵标准测试数据集，其中包括 4 种入侵类型，分别是：1) DOS，拒绝服务攻击，例如 ping-of-death, syn flood, smurf 等；2) PROBING，端口监视或扫描，例如

port-scan, ping-sweep 等；3) U2R，未授权的本地超级用户特权访问，例如 buffer overflow attacks；4) R2L，来自远程主机的未授权访问，例如 guessing password。由于数据集样本数量庞大，从中随机选取部分数据进行测试，具体见表 1。为使本文模型结果具有可比性，采用遗传算法优化回声状态神经网络(GA-ESN)、粒子群算法优化回声状态神经网络(PSO-ESN)进行对比实验。

**Table 1.** Sample set distribution
**表 1.** 样本集分布情况

| 入侵类型 | 训练样本 | 测试样本 |
|---|---|---|
| DoS | 2000 | 400 |
| Probe | 1000 | 200 |
| U2R | 500 | 100 |
| R2L | 500 | 100 |

## 3.2. 结果与分析

### 3.2.1. 检测结果对比

GA-ESN、PSO-ESN 和 BOA-ESN 的检测率和误报率的仿真结果如图 1 和图 2 所示。进行仔细分析

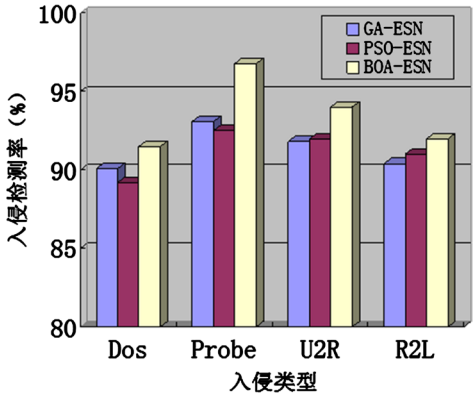


**Figure 1.** Comparison of intrusion detection rates for several models
**图 1.** 几种模型的入侵检测率比较

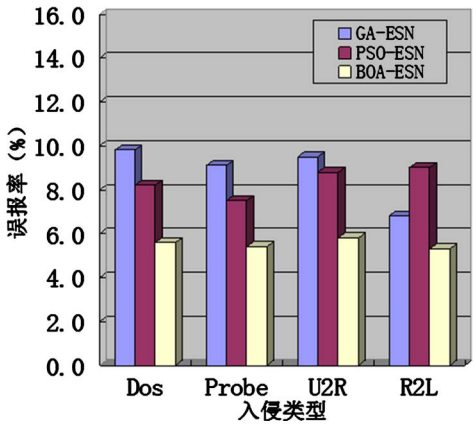


**Figure 2.** Comparison of false positive rates of several models
**图 2.** 几种模型的误报率比较

后可知，相对于模型 GA-ESN、PSO-ESN，蝙蝠优化算法优化的回声状态神经网络的入侵检测模型性能最佳，网络入侵检测的入侵检测率最高、误报率最低，具有明显的优势。

#### 3.2.2. 检测速度对比

为检测模型的运行效率，采用模型对测试集的检测时间作为指标，各模型的检测时间详见表 2。从表中数据可知，相对于 PSO-ESN、GA-ESN，BOA-ESN 检测速度更快，主要由于 BOA-ESN 采用了分层的技术，减少了计算时间，同时对 BA 算法进行了改进，加快了收敛速度，因此，BOA-ESN 能够更加满足现代网络入侵检测系统高效性、实时性的要求，具有广泛的应用前景。

**Table 2.** Comparison of detection time (s) of different models
**表 2.** 不同模型的检测时间(s)对比

| 入侵类型 | GA-ESN | PSO-ESN | BOA-ESN |
|---|---|---|---|
| DoS | 0.95 | 1.62 | 0.56 |
| Probe | 1.33 | 1.41 | 0.58 |
| U2R | 1.43 | 1.73 | 0.93 |
| R2L | 0.98 | 1.34 | 0.61 |

## 4. 结束语

针对回声状态神经网络参数优化的难题，比如参数的不当选择会导致算法收敛效果差等，提出了一种基于蝙蝠优化算法优化回声状态神经网络参数的入侵检测模型。该模型很好的将优化后的蝙蝠算法与回声状态神经网络进行了结合。在神经网络参数优化的过程中，我们既考虑了回声状态神经网络的训练误差反传，用误差改进对模型的训练，同时跟踪个体和群体的最佳权值来更新权值，这样，既充分利用了蝙蝠优化算法的全局搜索能力，又较好地利用了回声状态神经网络误差反向传播改进模型的特点，提高了模型的整体性能。仿真实验结果表明，相较于对比的两个模型，本文模型提高了网络入侵的检测率，同时降低了误报率，具有一定的实际应用价值和现实意义。

## 致 谢

首先，我要对我的授业导师李捷教授表示深深的感谢，感谢李老师在我研究生期间给我的方方面面的关心和照顾。李捷老师是一位非常和蔼近人的教授，他严于律己宽以待人，对工作认真积极，对科学一丝不苟的精神深深打动着我。李老师虽然每天工作都很辛苦，但是从不抱怨，而是乐观积极的去面对。他虽然忙，但是在指导和关心我的学习和生活上从未缺席，而且在研究上给了我很大的自由和及时的指导，谢谢您的支持和信任。

感谢于俊洋老师，于老师是一个聪明而又博学的人，在我的课题研究和论文写作方面都给了我很大的帮助和支持。还记得我第一次投稿的时候，心情一直有点焦虑浮躁，通常都是格式还没有改好，一些简单的语法和词序还没有修改好就着急找期刊投，这时候于老师总是很耐心的安慰和鼓励我，并能及时指出文章的各种不足。可以说没有于老师的认真指导和细心修改，我的第一篇论文是不会那么容易被期刊录用的。感谢于老师在每次我遇到问题时，总是认真细致的为我解答。

感谢在实验室里一起研究讨论的各位同学，感谢你们在科学研究以及撰写论文等方面给我的科学指导。和你们一起度过的时光是充实而快乐、紧张又幸福的，感谢你们的关心和帮助。

## 基金项目

本文由赛尔网络下一代互联网创新项目(NGII20160204)、网络与交换技术国家重点实验室开放课题资助项目(SKLNST-2016-2-23)资助。

## 参考文献

[1] 高阳. 机器学习在网络入侵检测中的应用浅谈[J]. 电脑迷, 2018(9): 178.

[2] 颜谦和, 颜珍平. 遗传算法优化的神经网络入侵检测系统[J]. 计算机仿真, 2011, 28(4): 141-144.

[3] 陈仕涛, 陈国龙, 郭文忠, 等. 基于粒子群优化和邻域约简的入侵检测日志数据特征选择[J]. 计算机研究与发展, 2010, 47(7): 1261-1267.

[4] Hong, S.-J., *et al*. (2011) A Novel Intrusion Detection System Based on Hierarchical Clustering and Support Vector Machines. *Expert Systems with Applications*, **38**, 306-313. https://doi.org/10.1016/j.eswa.2010.06.066

[5] 顾丽, 王广泽, 乔佩利. 基于改进遗传算法的入侵检测的研究[J]. 信息技术, 2009, 33(7): 58-61, 65.

[6] 杨富华, 彭刚. PCA-SVM 在网络入侵检测中的仿真研究[J]. 计算机仿真, 2011, 28(7): 146-149.

[7] Wang, X.S. (2010) A New Metaheuristic Bat-Inspired Algorithm. Nature Inspired Cooperative Strategies for Optimization, Studies in Computational Intelligence, Springer-Verlag, Berlin Heidelberg, 10: 65-74.

[8] Yang, X.S. (2011) Bat Algorithm for Multi-Objective Optimization. *International Journal of Bio-Inspired Computation* (*IJBIC*), **3**, 267-274. https://doi.org/10.1504/IJBIC.2011.042259

[9] Yang, X.S. and Gandomi, A.H. (2012) Bat Algorithm: A Novel Approach for Global Engineering Optimization. *Engineering Computation*, **29**, 267-289. https://doi.org/10.1108/02644401211235834

[10] 刘羿. 蝙蝠算法优化神经网络的网络入侵检测[J]. 计算机仿真, 2015, 32(2): 311-314.

[11] Denning, D.E. (2010) An Intrusion Detection Model. *IEEE Transaction on Software Engineering*, **13**, 222-232. https://doi.org/10.1109/TSE.1987.232894

[12] 李小剑, 谢晓尧. 基于支持向量机与 k 近邻相结合的网络入侵检测研究[J]. 贵州师范大学学报(自然科学版), 2015, 33(3): 86-91.

[13] 王云鹏, 张浩. 基于支持向量机的网络入侵检测算法综述[J]. 科学技术创新, 2017(25): 136-137.

[14] 谢康. 基于神经网络的入侵检测相关技术研究[D]: [博士学位论文]. 济南: 山东大学信息科学与工程学院, 2016.

[15] Scardapane, S. and Uncini, A. (2017) Semi-Supervised Echo State Networks for Audio Classification. *Cognitive Computation*, **9**, 125-135. https://doi.org/10.1007/s12559-016-9439-z

[16] 丁少华. 人工鱼算法优化神经网络的网络入侵检测[J]. 中国信息技术教育, 2014(20): 61.

[17] 贾会群, 魏仲慧, 何昕, 李沐雨. 基于神经网络与最小二乘法的车道线检测算法研究[J]. 汽车工程, 2018, 40(3): 363-368.

[18] 郭振洲, 王平, 马云峰, 王琦, 拱长青. 基于自适应权重和柯西变异的鲸鱼优化算法[J]. 微电子学与计算机, 2017, 34(9): 20-25.

[19] Wu, P.C. and Huang, K.C. (2006) Parallel Use of Multiplicative Congruential Random Number Generators. *Computer Physics Communications*, **175**, 25-29. https://doi.org/10.1016/j.cpc.2004.08.009

**知网检索的两种方式：**

1. 打开知网页面 http://kns.cnki.net/kns/brief/result.aspx?dbPrefix=WWJD
   下拉列表框选择：[ISSN]，输入期刊 ISSN：2161-8801，即可查询
2. 打开知网首页 http://cnki.net/
   左侧“国际文献总库”进入，输入文章标题，即可查询

投稿请点击：http://www.hanspub.org/Submission.aspx
期刊邮箱：csa@hanspub.org