Hindawi

无线通信和移动计算卷 2022，文章 ID 4252766，8 页 https:/ doi.org/10.1155/2022/4252766

研究文章

基于改进回声状态网络的网络流异常检测

陈明忠 ，1 邱斌，2 和 Jie Ji3

[1]汕头职业技术学院机电工程系，广东汕头 515078

[2]汕头职业技术学院计算机系，广东汕头 515078

[3]汕头大学网络与信息中心，广东汕头 515063

信件应寄给陈明忠;cmzgjjx@163.com 收到 2022 年 3 月 20 日;2022 年 6 月 23 日录用;发表于 2022 年 7 月 8 日学术编辑:Pierre-Martin Tardif

针对目前网络流量异常检测中存在的问题，设计了一种基于双环备用池回声状态网络的网络流量预测模型，解决了传统回声状态网络随机生成备用池的问题，提高了网络流量预测的准确性和实时性。然后，提出了一种基于动态阈值的异常检测方法，将预测值与真实值的差值作为判断异常发生的依据。仿真结果表明，改进的预测模型和异常检测方法能够有效检测网络流的异常行为，检测效果优于其他模型。

1.介绍

随着信息技术的不断发展，互联网上出现了越来越多的应用。目前，网络已经成为通信的主要载体。但网络攻击行为日益增多，网络安全问题也越来越突出。用户的网络行为体现在网络流量上，网络流量具有非线性、时效性、突然性、多样性[1]等特点。传统的线性预测方法已经不能适应现代网络发展的要求。近年来，神经网络、灰色模型、支持向量机、混合模型等非线性预测模型以良好的非线性映射和灵活的学习方法被广泛应用于场景预测。但目前仍存在预测精度低、预测速度慢[2]等问题。在此背景下，一种新型的递归神经网络——回声状态网络(echo state network，简称 ESN)应运而生。ESN 采用储备池结构，表现出优异的非线性处理能力和快速预测能力变形速度。但传统的 esn 使用单圈储备池，储备池的结构和权重是随机生成的。即使是具有相同参数的储备池也可能表现出显著的性能差异[3]。而且，训练的时间复杂度很高。这些因素将影响基于 ESN 网络流量的非线性表征和实时预测。

因此，为了保证网络流量预测的非线性和实时性，本文设计了一种基于双环储备池 ESN 的网络流量异常检测方法(ESN- dlrp)。首先，利用小波对网络流量进行归一化和去噪。其次，利用多环备用池 ESN 构建网络流量预测模型，解决了传统 ESN 随机产生的备用池问题，提高了网络流量预测的准确性和实时性;然后，我们可以得到网络流量的真值和预测值的关系图。最后，提出了一种基于动态阈值的异常检测方法，提取出真值与预测值之间偏差较大的报文。仿真结果表明，该模型

能有效检测异常网络流量行为，检测效果优于其他模型。

2.相关理论与技术研究

2.1.小波变换和小波去噪。小波变换是将三角函数基的无限长度变换为有限长度的衰减小波基。如图 1 所示。

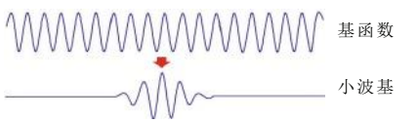

图 1:从基函数到小波基的变换。

因此，小波函数的表达式如下:

$$W_{a,\tau}=\frac{1}{\sqrt{a}}\int_{-\infty}^{+\infty}f(t)*\psi\left(\frac{t-\tau}{a}\right)dt. \qquad \text{ð1Þ}$$

该公式有两个变量:一是控制小波基伸缩的一个(尺度);另一个是控制小波基平移的 τ (平移)。对给定信号进行小波变换，就是将信号按照小波函数的小簇展开，即将信号表示为一系列不同尺度、不同平移的小波函数的线性组合。每一项的系数称为小波系数。

由于很多因素的影响，网络流量采集中可能会有一些噪声。噪声是一种无用的信号，它不仅浪费存储空间和传输时间，而且对网络流量预测造成干扰。一般来说，有用的网络流对应的是振幅较大的小波系数，而噪声对应的是振幅较小的小波系数[4]。所以我们选择一个阈值 λ 来与小波系数进行比较。如果小波系数小于 λ，可以认为是噪声，去噪[5]后必须去除才能得到网络流量。下面式(2)为阈值函数。

$$\hat{W}_{a,\tau}=\begin{cases}\operatorname{sgn}(W_{a,\tau})\times\left(|W_{a,\tau}|-\dfrac{2\lambda}{1+e^{|W_{a,\tau}|-\lambda}}\right), & |W_{a,\tau}|\geq\lambda,\\ 0, & |W_{a,\tau}|<\lambda.\end{cases} \qquad \text{ð2Þ}$$

*W α，τ 为原始小波系数。Ŵ α，τ 为去噪后的小波系数。*阈值函数是一个连续函数，因此当 jWa，τ j≥ λ 时，去噪后的网络流可以表示为:

$$f(x)=\operatorname{sgn}(x)\left(|x|-\frac{2\lambda}{1+e^{|x|-\lambda}}\right). \qquad \text{ð3Þ}$$

2.2.传统回声状态网络

2.2.1.传统回声状态网络的结构。回声状态网络(Echo state network，简称 ESN)是一种由输入层、储备池和输出层组成的新型递归神经网络。作为 ESN 的核心，储备池结构随机生成，包含大规模稀疏连接的神经元。这些神经元包含了系统的运行状态，具有很强的非线性学习能力，可以根据已知值预测未来的值

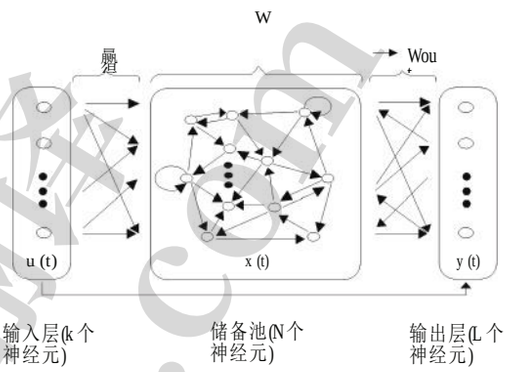

图 2:传统 ESN 的基本结构

的网络流量时间序列[6].;图 2 展示了传统 ESN 的基本结构。

假设输入层有 K 个神经元，储备池有 N 个神经元，输出层有 L 个神经元，当时间为 t 时，输入层、储备池和输出层构造的信号分别为如下方程:

$$u(t)=[u_1(t),u_2(t),\cdots,u_K(t)]T, \qquad \text{ð4Þ}$$

$$y(t)=[y_1(t),y_2(t),\cdots,y_L(t)]T. \qquad \text{ð5Þ ð6Þ}$$

如果输入层与储备层之间的连接矩阵为 Win ðN × KÞ，那么储备层内部的连接矩阵为 W ðN × NÞ，储备层与输出层之间的连接矩阵为 WoutðL × NÞ。若当前时刻的输入为 u ðt + 1Þ，则储备池的状态方程为 ð7Þ

$$x(t+1)=f_{\text{in}}(W_{\text{in}}\times u(t+1)+Wx(t)). \qquad \text{ð8Þ}$$

输出层的状态方程为 Eq(7)中，f in是储备池的激励函数。*X ðt + 1Þ为储备池在当前时刻的状态。X ðtÞ 是储备池最后时刻的状态。*在式(8)中，f out为输出层的激励函数，y ðt + 1Þ为当前时刻的输出状态。

$$y(t+1)=f_{\text{out}}(W_{\text{out}}\times x(t+1))$$

2.2.2.ESN 的培训。将 ESN 网络流数据分为训练样本和验证样本。通过输入连接矩阵 Win，训练样本 fu ðt + 1Þ， y ðt + 1Þg 进入储备池。在方程式。式(7)和式(8)中，Win 和 W在训练前随机生成。唯一需要训练的因素是 Wout，它

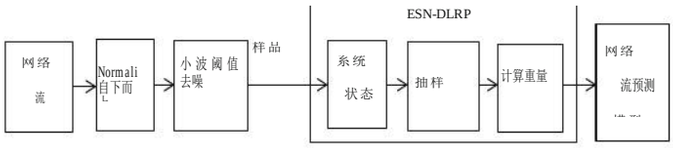

图 3:网络流量模型预测的构建过程。

表示 ESN 的训练过程就是确定 Wout[7]的过程。在 Eq.(7)中，由于 Win 和 W 是随机生成的，我们可以立即得到 x ðt + 1Þ。在式(8)中，Wout 可以用 x ðt + 1Þ 和 y ðt + 1Þ 来计算。

与传统神经网络相比，ESN 网络具有更强的网络流量预测能力。但采用单环储备池，储备池结构和权重随机生成，在一定程度上影响了 ESN 对网络流量的非线性表征和实时预测[8]。

3.基于改进 ESN 的网络流量预测模型设计

为保证网络流量预测的非线性和实时性，本文设计了一种双环储备池结构的 ESN 网络流量预测方法。对某 Web 系统采集原始网络流量后，进行归一化和去噪，得到去噪后的网络流量数据。然后，利用双环储备池的 ESN 号构建网络流量预测模型;这样的过程如图 3 所示。

3.1.网络流数据的预处理

(1)对采集到的原始网络流量进行归一化处理，得到[0.1,1]之间的网络流量数据。设 xmin 和 xmax 分别表示网络流量的最小值和最大值，x '表示归一化后的数据

$$x' = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \times 0.8 + 0.1. \qquad (9)$$

(2)采用小波阈值法对归一化网络流进行去噪，消除噪声，避免对 normal 信号的干扰

(3)将去噪后的网络流分为训练样本和验证样本，将训练样本输入 ESN-DLRP 中进行学习

3.2.ESN-DLRP 的设计与预测方法

3.2.1.ESN-DLRP 的设计。基于双环储备池的回声状态网络(Echo state network based on 双环 reserve pool，简称 ESN-DLRP)如图 4 所示。储备池中的每个神经元都以圆形的方式与相邻神经元连接，形成第一个环。然后，以第一个神经元 A为起点进行欺骗

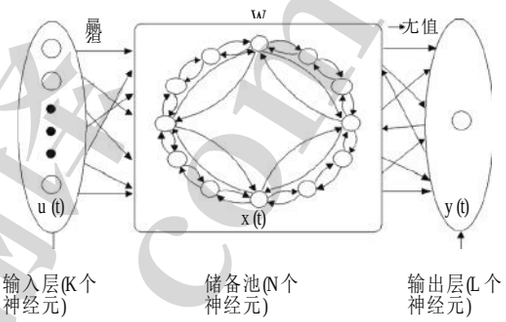

图 4:ESN-DLRP 的结构

在 d 区间(本文中 d = 3)与神经元 B 进行循环 nect，然后 B 与后续 d 区间的神经元 C 连接，以此形成第二个环，从而构建 ESN-DLRP-d (d 为第二个环的神经元区间)。这种双环备用池结构避免了传统 ESN 产生备用池的随机性，增强了备用池中神经元之间的连通性，提高了对网络流[9]的非线性表征能力。

ESN-DLRP 的预测原理是利用输入层在时刻 t 的网络流序列来预测输出层在时刻 t + h 的网络流数据 tr ðt + hÞ。在时刻 t，输入层的输入向量是 u ðtÞ =½tr ðt− KÞ， tr ðt - K + 1Þ，⋯ ，tr ðt− 1Þ， tr ðtÞ?^{T}，输出层的输出向量为 y ðtÞ =½tr ðt + hÞ?^{T}。*H 为预测长度。*所以在 t 时刻，储备池的内部状态向量为

$$x(t) = f_{\text{in}}(W_{\text{in}}u(t) + Wx(t-1)). \qquad (10)$$

在 Eq.(10)中，f In 为内部 reserve pool 中的激励函数。*X ðt− 1Þ 为储备池在前一时刻的状态。*t 时刻，输出层的状态向量为

$$tr(t+h) = y(t) = W_{\text{out}}x(t). \qquad (11)$$

在 Eq.(11)中，f out 是输出层的激励函数。

3.2.2.ESN-DLRP 的训练和预测模型的构建

(1)在储备池中建立连接矩阵 W。给定神经元 i, i = 1,1 + d, 1 + 2d，⋯ ，1 + ðN/d− 2Þ • d.设置储备池连接的元素

matrix to $W_{i,i+d} = r$, $W_{i+d,i} = r$. When $i = 1 + (N/d - 2) \cdot d$, set elements $W_{i,1} = r$, $W_{1,i} = r$. The weight

值 r∈ð0, 1Þ， N为神经元的编号

(2)根据 ESN- DLRP 的预测特性，训练样本的输入-输出对 futrain ðtÞ，

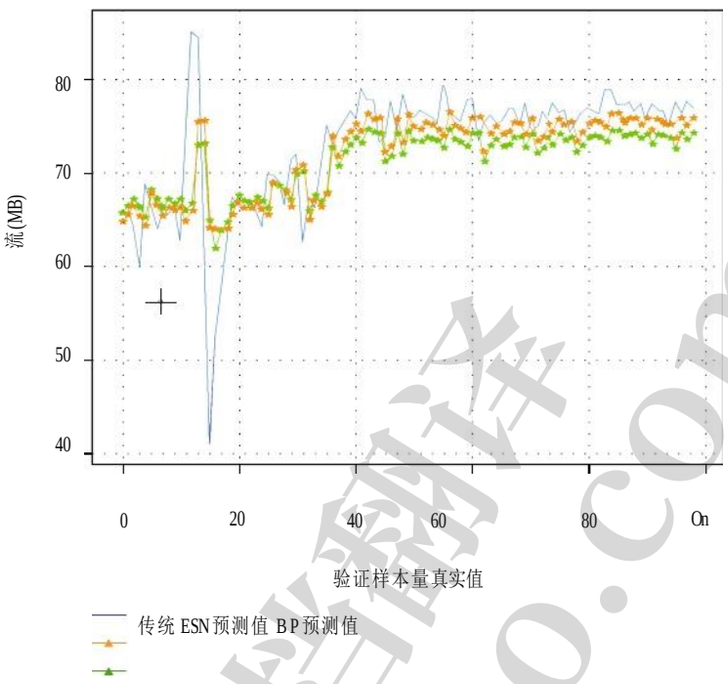

图 5:BP 与传统 ESN 预测效果对比

通过输入连接矩阵 Win[10]，建立 ytrain ðt + hÞ， t = 1,2，⋯，Tg 并进入储备池

(3)抽样阶段(The Sampling Stage)。开始收集储备池中的状态向量 x ðtÞ 以及从时间 t0 到时间 T结束的相关样本数据 ytrain ðt + hÞ，从而建立状态矩阵 x ðN, T− t0 + 1Þ =½x ðt0Þ， x ðt0 + 1Þ，⋯，x ðTÞ?和预期输出

$$\text{matrix}\quad Y(L, T-t_0+1)=[y_{\text{train}}(t_0+h), y_{\text{train}}(t_0+1+h), \cdots, y_{\text{train}}(T+h)]$$

(4)权重计算。根据状态矩阵 X 和期望输出矩阵 Y 得到输出权值

$$W_{\text{out}}, \text{ that is, } Y \approx W_{\text{out}}X, \text{ because the actual output matrix } \hat{Y}=[y(t_0+h), y(t_0+1+h), \cdots, y(T+h)] \text{ is}$$

与状态矩阵 X 成线性关系，即 Ŷ = W_{out}*X.训练目的是通过实际输入矩阵 Ŷ 逼近预期输出矩阵 Y，得到* Y = W_{out}*X.然后，* W_{out} = X + Y· 。其中 X+是 X的伪逆矩阵

(5)将计算无值放入等式。(10)和式(11)得到网络流量的预测模式

3.2.3.预测验证样本的输出。根据网络流的预测模型，预测验证样本的输出，建立输入-输出对[11]。

4.预测模型的预测试验与性能评价

本文用 MATLAB 中的商业数学软件对一个 Web 应用系统，每 5 分钟对网络流数据进行一次采样，分别取 100 个样本，输入 4 种模型:BP(back propagation)神经网络，tradi-定义 ESN、ESN- dlrp -3 和 ESN- dlrp -8，记录未来一段时间的预测值，并将预测值与实际值进行比较。四种模型的预测结果如图 5、6、7 所示。

下面用均方根误差(RMSE)和平均绝对百分比误差(MAPE)[12]来分析四个模型的预测性能。

(1) RMSE。时间序列数据的预测值与真实值之间偏差的平方和。将这个和除以样本数量。得到商的平方根。可以反映预测模型的预测精度。数值越小，预测就越准确

$$\text{RMSE}=\sqrt{\frac{\sum_{i=1}^{n}(y_i-\hat{y}_i)^2}{n}}. \qquad \text{ð12Þ}$$

(2)日军。将时间序列数据的预测值与真实值相减，得到差值。用差值除以真实值。得到商的绝对值的平均值。平均值越小，预测就越准确

$$\text{MAPE}=\frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i-\hat{y}_i}{\hat{y}_i}\right|\times 100\%. \qquad \text{ð13Þ}$$

Yi 是 predictive value, Yi 是 real value, n 是 sample number。

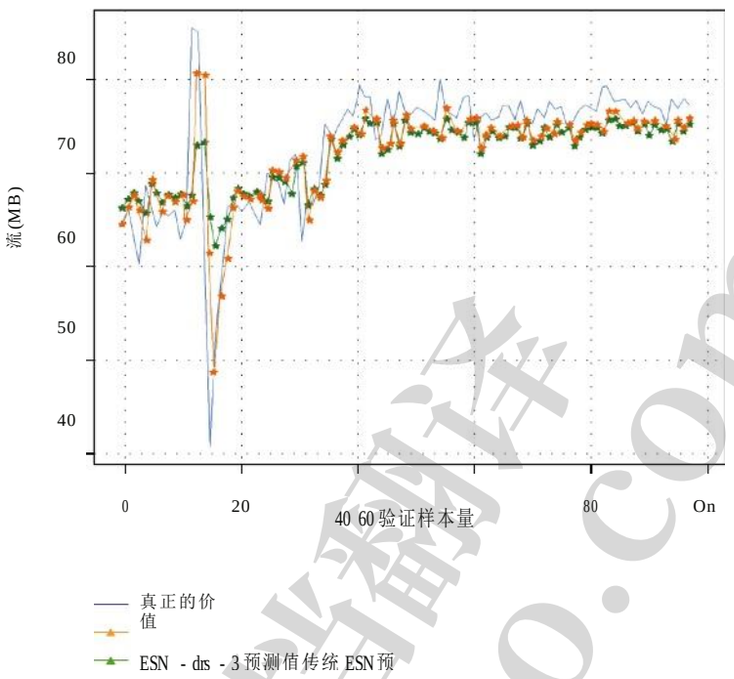

图 6:传统 ESN 与 ESN-DLRP 预测效果对比 90

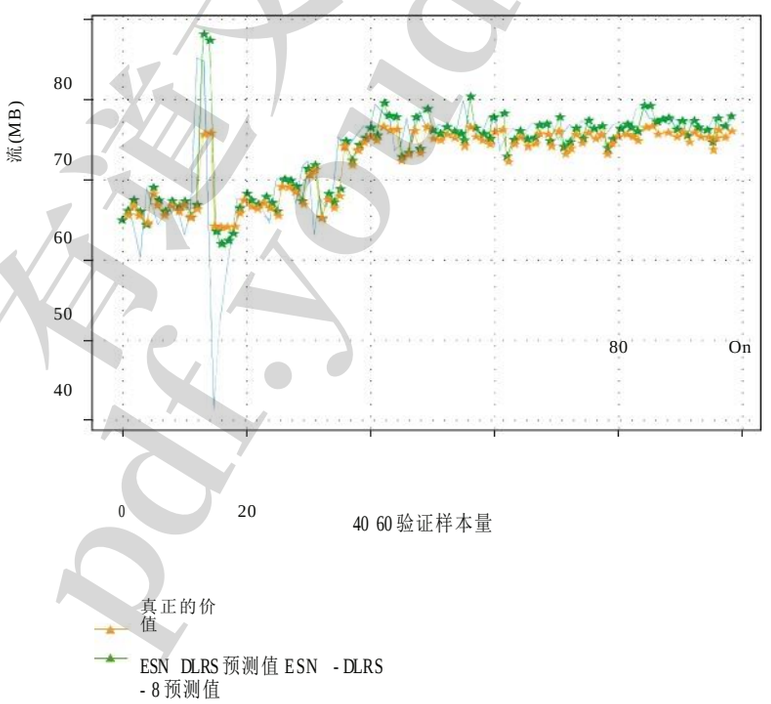

图 7:ESN-DLRP-3 与 ESN-DLRP-8 预测效果比较

根据图 4 -图 6 的数据和测试时间，四种模型的预测性能如表 1 所示。在预测误差方面，传统 ESN 网络小于 BP 神经网络，ESN- DLRP 明显小于 BP 神经网络和传统 ESN 网络。在双环的情况下，第二环内神经元间距越大，预测误差越小，但差异不显著。即 ESN-DLRP-8 的预测误差略小于 ESN-DLRP-3。从 100 个样本的测试时间来看，ESN-DLRP 明显小于 BP 神经

网络和传统 ESN。在双环情况下，ESN-DLRP-8 的测试时间略短于 ESN-DLRP-3。也就是说，使用 ESN-DLRP-3 模型可以获得更好的预测精度和测试率。

5.异常网络流量检测

在 Web 系统正常运行时，利用 ESN-DLRP 可以对网络流量进行准确预测，即网络流量的预测值与真实值接近，但当系统发生异常事件时，会出现较大偏差

表 1:不同模型网络流量预测性能比较。

预测模型	均方根误差(%)	日军(%)	测试时间(ms)
BP 神经网络	5.19	5.60	45.23
传统 ESN	5.01	5.30	30.29
ESN-DLRP-3	4.65	4.11	22.95
ESN-DLRP-8	4.57	3.98	20.34

表 2:不同阈值下的检测效果表。

序列号	阈值	检出率	误报率
1	10%	97%	39%
2	20%	95%	34%
3.	30%	92%	24%
4	40%	90%	22%
5	50%	88%	8.6%
6	60%	87%	7.0%
7	70%	86%	2.8%
8	80%	85%	0.7%
9	90%	80%	0.5%
10	100%	76%	0.2%

预测值与实际值[13]之间。本文采用滑动窗口机制，每次发送或接收的数据帧数称为窗长。根据窗口内网络流量的真实值与预测值的差值来判断数据点是否异常。

如果当前时间为 t，真实值为 yðtÞ，预测值为 yðtÞ，滑动窗口长度为 l，我们可以根据绝对偏差值 D ðtÞ，平均偏差值 u，相对偏差值 P ðtÞ 来估计真实值与预测值的差值。

(1)绝对偏差值 D ðtÞ。数据点的真实值与预测值之差的绝对值

$$D(t) = |\hat{y}(t) - y(t)|. \quad (14)$$

(2)平均偏差值 u，滑动阶段绝对偏差值的平均值

$$u = \frac{1}{l}\sum_{t=0}^{l} D(t). \quad (15)$$

(3)相对偏差值 P ðtÞ。数据点的绝对偏差值与平均偏差值的比例。然后比例值乘以 100%。

$$P(t) = \frac{|\hat{y}(t) - y(t)|}{u} \times 100\%. \quad (16)$$

Anomaly 本质上是数据点偏离 normal 值的行为。在本文中，每个滑动 win- dow 的平均值 u 是正常值的标准。如果数据点的绝对偏差值(t)不高于平均值 u，则认为是正态[14]。如果高于 u，则需要进一步估计是否高于阈值 K，如果不高于 K 则认为正常，否则为异常。每个滑动窗口的异常检测过程如下:

步骤 1。获得滑动窗口中每个数据点的实值和预测值对。

步骤 2。计算每个数据点的绝对偏差值 D ðtÞ。

步骤 3。统计滑动区间内数据序列的平均偏差值 u。

步骤 4。将数据点的绝对偏差值 D ðtÞ 与 u进行比较，如果 D ðtÞ≤u，则数据正常;否则，则需要执行步骤 5。

第 5 步。确定动态阈值。

选择一个滑动范围内的数据作为检测样本集，记录样本集中正常数据、异常数据和总数据的个数。异常检测的效果可以通过检测准确率(检出率)和误报率两个指标来衡量。检出率是指

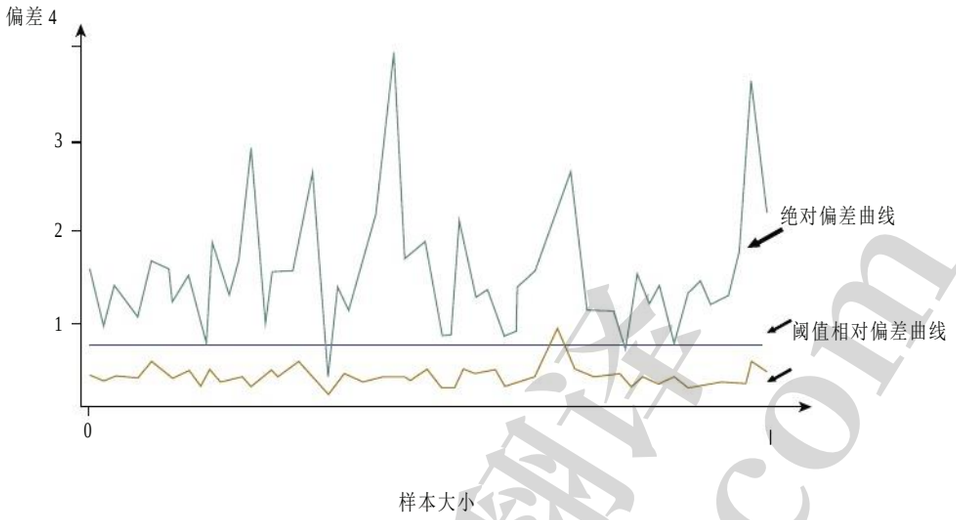

图 8:绝对偏差和相对偏差值曲线。

到检测到的异常量中正确量与总量的比值。误报率是指错误检测次数与总检测次数[15]之比。选取不同阈值时，该样本集中数据的检出率和误报率如表 2 所示。

表 2 显示，当阈值为 80%时，该检测方法检出率高，误报率低，因此本文确定异常检测的阈值为 80%。图 8 显示了样本集中各数据点的绝对偏差和相对偏差曲线。从图中可以看出，使用相对偏差解决了绝对偏差在峰值和谷值处容易产生误检的问题。通过相对偏差值与阈值的比较，可以准确检测出 Web 系统中的异常数据。

步骤 6。我们计算了每个数据点的相对偏差 P ðtÞ，并与阈值 K进行了比较。如果 P ðtÞ≤K，则数据正常;否则为异常。

6.结论

网络流量异常检测是网络安全领域的一个热点。本文主要从三个方面进行研究:首先，分析了传统 ESN 在网络流量预测中面临的问题;其次，设计了基于 ESN 的网络流量预测模型——DLRP。通过实验，比较了神经网络、传统 ESN、ESN- dlrp -3 和 ESN- dlrp -8 模型的性能。结果表明，ESN-DLRP- 3 模型具有良好的预测精度和测试率。第三，提出了一种基于动态阈值的异常检测方法。阈值是动态确定的。如果数据点真实值与预测值之间的相对偏差超过阈值，则将该数据判定为异常数据。

数据可用性

在当前研究期间生成和/或分析的数据集可根据合理要求从通信作者处获得。

利益冲突

作者声明，就目前的研究而言，他们没有任何利益冲突需要报告。

致谢

本文由广东省高校特色创新项目(no. 1)资助。2020 ktscx305)。

参考文献

宋杰，樊平，高波，周旭，李振，“基于深度学习的网络流量预测研究”，计算机工程与应用，vol. 57, no. 6。10，第 1-9 页，2021。

[2] W. Jian，网络异常检测关键技术研究，南京邮电大学，2020。

郑 d .，马丽玲，“基于小波分析的异常网络流量检测”，计算机科学与技术，vol. 46, no. 7。第 8 期，第 178-182 页，2019。

张志达，“基于小波系数感知的网络流量预测系统”，《科技与电子信息》，第 1 期。第 1 期，131-135 页，2019 年。

林明柱，“基于小波去噪和回声状态网络的光电检测系统状态识别”，《激光》杂志，第 42 卷，第 1 期。第 5 期，143-146 页，2021 年。

魏勇，张杰，“基于时间序列分析的异常网络流量检测”，吉林大学学报(自然科学版)，第 55 卷，第 6 期。第 5 期，第 1249-1254 页，2017。

闫磊，“基于时间序列分析的异常网络流量检测”，现代电子技术，第40卷，第1期。第7期，85-87页，2017年。

程磊，基于回声神经网络时间序列预测系统的研究与实现，江苏大学，2020。

宋宗江，戚胜，尚明，沈丹，“基于多尺度回声状态网络的微电网等效建模”，计量学报，第42卷，第1期。第7期，第923-929页，2021。

Y. Xinyan，基于高级 ESN 的网络流量预测研究，南京邮电大学，2019。

[11] S.铁宁，回声状态网络优化与应用研究，广西师范大学，2019。

任仁军，王 w，“基于深度神经网络的网络流量预测模型仿真”，计算机仿真，第38卷，第6期。第6期，第475-479页，2021。

黄勇，孙旭，“基于深度回声状态网络的网络流量预测模型”，《南京邮电大学学报》(自然科学版)，第38卷第1期。第5期，85-90页，2018年。

杨泽勇，网络流量识别与预测系统关键技术研究，西安电子科技大学，2018。

张慧，罗赟，王旭，华建华，王 w，“基于 KL 长度的自适应阈值网络流量异常检测”，计算机工程，vol. 45, no. 7。第4期，108 - 113页，2019年。