研究与开发

基于深度学习的实时 DDoS 攻击检测

李传煌[1]，孙正君[1]，袁小雍[2]，李晓林[2]，龚梁[1]，王伟明[1]
（1. 浙江工商大学信息与电子工程学院，浙江 杭州 310018；
2. 美国佛罗里达大学大规模智能系统实验室，美国 佛罗里达州 盖恩斯维尔 32611）

摘　要：分布式拒绝服务（DDoS）攻击是一种分布式、协作式的大规模网络攻击方式，提出了一种基于深度学习的 DDoS 攻击检测方法，该方法包含特征处理和模型检测两个阶段：特征处理阶段对输入的数据分组进行特征提取、格式转换和维度重构；模型检测阶段将处理后的特征输入深度学习网络模型进行检测，判断输入的数据分组是否为 DDoS 攻击分组。通过 ISCX2012 数据集训练模型，并通过实时的 DDoS 攻击对模型进行验证。结果表明，基于深度学习的 DDoS 攻击检测方法具有高检测精度、对软硬件设备依赖小、深度学习网络模型易于更新等优点。

关键词：分布式拒绝服务；拒绝服务；深度学习
中图分类号：TP393　　**文献标识码：**A
doi: 10.11959/j.issn.1000−0801.2017191

Real-time DDoS attack detection based on deep learning

LI Chuanhuang[1], SUN Zhengjun[1], YUAN Xiaoyong[2], LI Xiaolin[2], GONG Liang[1],WANG Weiming[1]
1. Institute of Communication and Engineering, Zhejiang Gongshang University, Hangzhou 310018, China
2. Li Lab, University of Florida, Gainesville, Florida 32611, USA

Abstract: Distributed denial of service (DDoS) is a special form of denial of service (DoS) attack based on denial of service(DoS). It is a distributed, collaborative large-scale network attack. A DDoS detection method based on deep learning was presented. The method included two stages: feature processing and model detection: feature extraction, format conversion and dimension reconstruction of the input data packet was performed in feature processing stage; in the model detection stage, the processed features were input to the depth learning network model to detect whether the input data packets was DDoS attack packet. The model was trained by the ISCX2012 dataset, and the model was validated by real-time DDoS attack. The experimental results show that DDoS attack detection method based on deep learning has high detection precision, little dependency on hardware and software equipment, and the model of depth learning network is easy to update.

Key words: distributed denial of service, denial of service, deep learning

收稿日期：2017−03−14；修回日期：2017−06−07
基金项目：国家高技术研究发展计划（“863”计划）基金资助项目（No.2015AA011901）；国家自然科学基金资助项目（No.61402408，No.61379120）；浙江省重点研发计划基金资助项目（No.2017C03058）
Foundation Items: The National High Technology Research and Development Program (863 Program)(No.2015AA011901), The National Natural Science Foundation of China(No.61402408, No.61379120), Zhejiang's Key Project of Research and Development Plan(No.2017C03058)

1 引言

DoS（denial of service，拒绝服务）攻击是当今网络环境中最常见的攻击手段之一，其基本原理是通过采取任意攻击方式，最终致使被攻击目标的系统资源耗尽甚至崩溃，被攻击目标“拒绝”为正常接入用户提供所需服务。DDoS（distributed denial of service，分布式拒绝服务）攻击是一种分布式、大范围协同作战的危害性更强的网络攻击方式，攻击者利用其所控制的数目众多的傀儡机，同时向被攻击目标发起 DoS 攻击，以若干倍于“一对一”DoS 攻击方式的攻击规模对被攻击目标发起网络攻击。

攻击检测是主要的 DDoS 防御机制之一，然而因为在大多数情况下攻击流量非常类似于合法流量，导致 DDoS 攻击很难实现自动检测，攻击者时常利用这一特性发起 DDoS 攻击，在多数情况下，低速率小流量的攻击活动在非常早的阶段常被误认为是合法活动[1]。许多研究人员尝试采用统计机器学习方法来检测 DDoS 攻击，虽然统计机器学习方法是基于统计特征对 DDoS 攻击进行分类，性能优于统计方法，但这种方法仍存在一些缺点，如需要丰富的网络专业知识和经验；在 DDoS 检测过程中需要复杂的人工提取适当的统计特征；仅限于一个或几个 DDoS 攻击向量；需要更新其模型和阈值以满足系统和攻击矢量的变化；对于低速率的 DDoS 攻击检测效果较差等[2]。

本文提出了一种基于深度学习的 DDoS 检测方法，对上述问题和不足进行分析改进，该方法包含特征处理和模型检测两个阶段：特征处理阶段对输入的数据分组进行特征提取、格式转换和维度重构，简化了机器学习方法中复杂繁琐的特征选择的过程，同时，得到多个 DDoS 攻击向量，解决了攻击向量单一性问题；模型检测阶段将处理后的特征输入深度学习网络模型进行检测，判断输入的数据分组是否为 DDoS 攻击分组。本文使用大规模数据集训练深度学习模型，基于深度学习的DDoS检测方法利用不同的神经网络模型：卷积神经网络（convolutional neural network，CNN）[3]、长期短期记忆神经网络（LSTM）[4]、门控循环单元（GRU）[5]，可以根据攻击状况自动调节模型权重，并且有效解决了低速率 DDoS 攻击检测精度低的问题。在实验中，处理来自 ISCX2012（Information Security Centre of Excellence 2012）[6]数据集的两天的网络流量，并对机器学习中随机森林方法和本文的深度学习方法在检测精度等方面进行对比，最后，通过实时的 DDoS 攻击对模型进行验证。

2 相关工作

当前国内外针对 DDoS 攻击的防御技术的研究，大多基于网络入侵检测[7]的方法进行实现，但随着 DDoS 攻击技术的发展，现有的 DDoS 攻击防御研究现状形势严峻，在网络流量处理方面，新型 DDoS 攻击检测等方面存在较高误报率和漏报率。

参考文献[8]采用 CAT（change-aggregation tree）机制对流经同一个 ISP 网络中的路由器流量进行协同分析，对路由器每个接口的流量分布情况进行分析，以发现异常流量报警。参考文献[9]提出通过采用交叉相关性和权值向量分析骨干网节点流量，检测恒速流量、增速流量等种类型的 DDoS 攻击的方法。这种检测方式不能够有效地区分 DDoS 攻击和大流量访问，误报率较高。

参考文献[10]通过比较 4 个重要信息熵测量（Hartley entropy、Shannon entropy、Renyin++s entropy、Renyin++s generalized entropy）检测 DDoS 攻击，根据这 4 个信息熵，计算网络中的信息距离并检测低速率和高速率 DDoS 攻击。参考文献[11]通过两次统计 t 检验来识别 DDoS 攻击 SYN 到达速率和 SYN、ACK 分组的数目。

然而大多数统计方法虽然在特定 DDoS 攻击方法中表现良好，但需要在特征选择方面进行大量繁杂的预处理工作。

参考文献[12]提出一种基于随机森林分类模型的 DDoS 检测方法，将数据流信息熵作为分类标准，对 3 种常见的 DDoS 攻击方式进行特征分析，在此基础上使用基于随机森林分类模型分别对 3 类 DDoS 攻击方式进行分类检测，该模型能够较为准确地区分正常流量和攻击流量。但是这种方式只适用于针对单个分组的检测，对于多重数据分组检测准确率较低。

3　基于深度学习的 DDoS 攻击检测方法

3.1　深度学习模型及算法

卷积神经网络是一种前馈人工神经网络，是由 Lecun[13]等人最早提出的，当前已是语音和图像识别领域的热门应用及研究问题。卷积神经网络主要利用了 3 个基本思想：局部感受野、权值共享和池化。局部感受野使每个神经元映射到局部特征，从而减少需要训练的权值参数；权值共享保证了同一个卷积核中所有神经元的权值都相同，因此能够大大减少网络中的训练参数；通过池化能够减小特征的规模，并且能够确保特征的不变性。因此，通过卷积神经网络的能够保证输入特征在位移、倾斜、比例缩放或其他变形后的顽健性，卷积神经网络模型如图 1 所示。

假设在卷积层之后有一个 $N\times N$ 大小的神经元层，如果用一个 $m\times m$ 的滤波器 w，则卷积层的输出大小为$(N-m+1)\times(N-m+1)$。为了计算每层中神经元的输入 x_{ij}^{l}，就需要对前一层的神经元中滤波器的权重求和，x_{ij}^{l} 的计算式为：

$$x_{ij}^{l}=\sum_{a=0}^{m-1}\sum_{b=0}^{m-1}w_{ab}y_{(i+a)(j+b)}^{l-1} \tag{1}$$

其中，用到了非线性计算式：

$$y_{ij}^{l}=\sigma(x_{ij}^{l}) \tag{2}$$

假设卷积层的误差函数为 E，并知道当前卷积层的误差值，则可以求出前一层的误差值和卷积层的权重梯度。因此，根据当前层的误差值 $\frac{\partial E}{\partial y_{ij}^{l}}$，可求出前一层误差值。通过应用链式法则可求出权重梯度$\left(\frac{\partial E}{\partial w_{ab}}\right)$，计算式如下：

$$\frac{\partial E}{\partial w_{ab}}=\sum_{i=0}^{N-m}\sum_{j=0}^{N-m}\frac{\partial E}{\partial x_{ij}^{l}}\frac{\partial x_{ij}^{l}}{\partial w_{ab}}=\sum_{i=0}^{N-m}\sum_{j=0}^{N-m}\frac{\partial E}{\partial x_{ij}^{l}}y_{(i+a)(i+b)}^{l-1} \tag{3}$$

根据式（3）可知，$y_{(i+a)(i+b)}^{l-1}=\frac{\partial x_{ij}^{l}}{\partial w_{ab}}$。其中，$\frac{\partial E}{\partial x_{ij}^{l}}$ 的计算式如下：

$$\frac{\partial E}{\partial x_{ij}^{l}}=\frac{\partial E}{\partial y_{ij}^{l}}\frac{\partial y_{ij}^{l}}{\partial x_{ij}^{l}}=\frac{\partial E}{\partial y_{ij}^{l}}\frac{\partial}{\partial x_{ij}^{l}}(\sigma(x_{ij}^{l}))=\frac{\partial E}{\partial y_{ij}^{l}}\sigma'(x_{ij}^{l}) \tag{4}$$

为了计算卷积层的权重，则需要将误差反向传递到前一层，而：

$$\frac{\partial E}{\partial y_{ij}^{l-1}}=\sum_{a=0}^{m-1}\sum_{b=0}^{m-1}\frac{\partial E}{\partial x_{(i-a)(j-b)}^{l}}\frac{\partial x_{(i-a)(j-b)}^{l}}{\partial y_{ij}^{l-1}}=\sum_{a=0}^{m-1}\sum_{b=0}^{m-1}\frac{\partial E}{\partial x_{(i-a)(j-b)}^{l}}w_{ab} \tag{5}$$

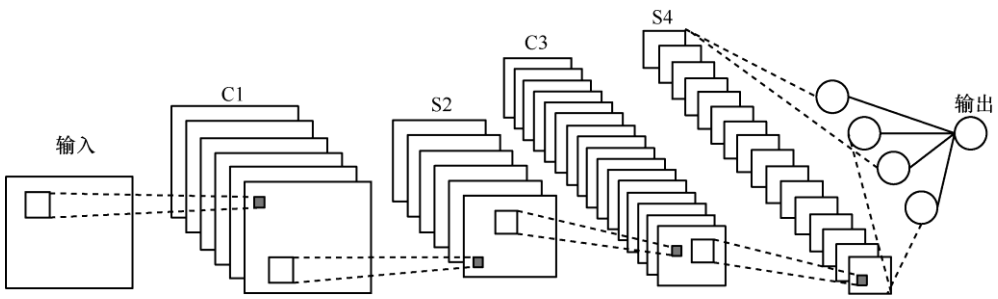


图 1　卷积神经网络模型

从而可得出 $w_{ab}=\frac{\partial x_{(i-a)(j-b)}^{l}}{\partial y_{ij}^{l-1}}$。

卷积神经网络是关于时间深度的神经网络，而递归神经网络（recurrent neural network，RNN）是关于空间深度的神经网络。在 1990 年，递归神经网络的概念就被 Kamijo 等人[14]提出。在前馈神经网络中，信息是按照从输入层到输出层的方向定向传递的，而递归神经网络打破了前馈神经网络中定向传递信息的限制。递归神经网络不同于前馈神经网络：前馈神经网络从输入层到输出层，层与层之间都是前后全连接的，每层内部的神经元之间没有连接关系；而递归神经网络的隐藏层内部的神经元之间是有连接关系的，在某一时刻该层中神经元的输入包括了输入层神经元的输入、同一层内其他神经元的输入和前一时刻神经元自身的输出。递归神经网络模型如图 2 所示。

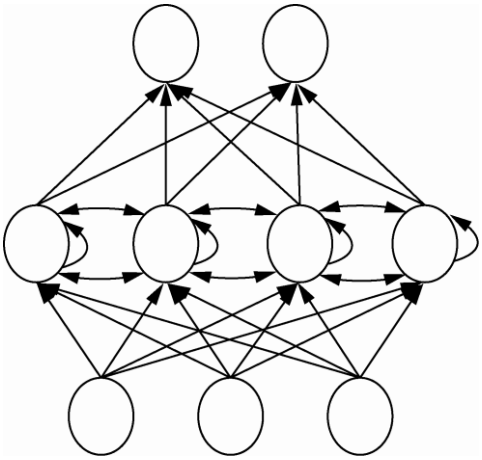


图 2　递归神经网络模型

对于一个简单的递归神经网络模型，隐藏层内部以递归的方式传递信息。不同于传统的神经网络，递归神经网络的隐藏层中的每一层都共享参数 U、V、W，因此能够大大减少该网络所需学习的参数。隐藏层内部神经元的向前传递过程如下：

$$a_j(t)=\sum_i w_{ij}x_i(t)+\sum_k u_{jk}s_k(t-1) \quad (6)$$

$$s_j(t)=f(a_j(t)) \quad (7)$$

$$o_q(t)=\sum_j v_{jq}s_j(t) \quad (8)$$

其中，函数 f 为神经元的激活函数，一般为非线性函数，w_{ij} 为输入特征与隐藏层中神经元之间的权重，u_{jk} 为隐藏层中神经元与前一层神经元之间的权重，v_{jq} 为隐藏层中神经元与输出特征之间的权重。隐藏层的误差不仅与当前时刻的输出 $o(t)$有关，还有 t+1 时刻该层传递回来的 $s(t+1)$有关。

Hochreiter S[15]在 1997 年提出了一种改进的递归神经网络模型长短期记忆（long short-term memory，LSTM）模型。他们认为传统的递归神经网络模型把隐藏层作为模型的记忆模块，与模型中的其他部分存在直接连接的关系，这才导致了梯度消失和梯度爆炸问题的产生。之后，多位研究者重新设计了传统递归神经网络的记忆模块，在神经元内添加了输入门（input gate）、输出门（output gate）和遗忘门（forget gate）这 3 个模块来加强整个 LSTM 网络的记忆能力，从而能够有效解决梯度消失和梯度爆炸问题，LSTM 中神经元的结构如图 3 所示。

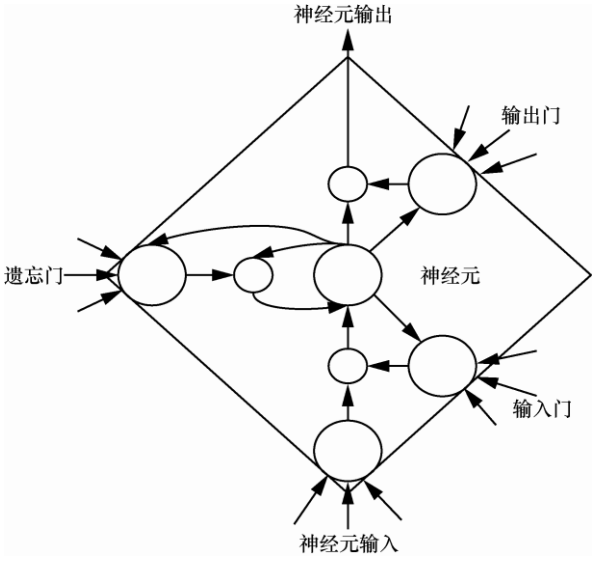


图 3　LSTM 网络中神经元的结构

3.2　特征处理阶段

特征处理阶段的处理流程如图 4 所示。

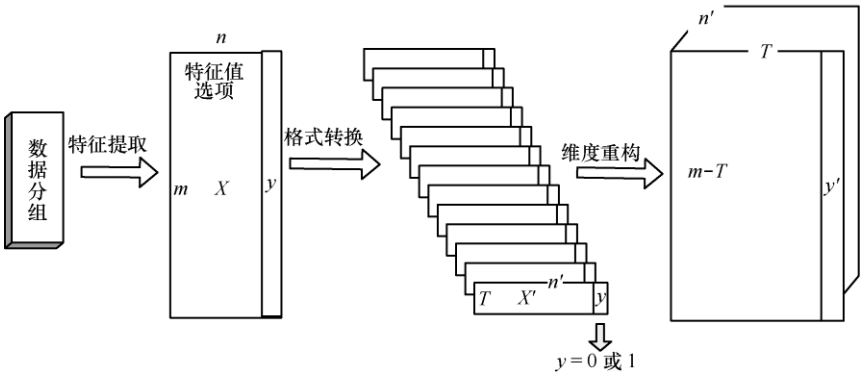


图 4　特征处理阶段的处理流程

对输入的 m 个数据分组分别提取 20 个报文字段作为特征值字段，并将这 20 个特征值字段划分为文本类型字段、数值类型字段和布尔类型字段 3 种类型，表 1 为提取的特征字段。

表 1　网络数据报的 20 个特征字段及实例值

特征值字段	数据类型	字段实例
frame.encap type	布尔类型	1
frame.len	数值类型	805
frame.protocols	文本类型	eth：ip：tcp：http:data
http.time	数值类型	0
icmp.length	数值类型	203
icmp.type	布尔类型	3
irc.request.command	文本类型	pong
irc.response.command	文本类型	462，join，353，366
tcp.ack	数值类型	2.692×10^3
tcp.analysis.ack_rtt	数值类型	0
tcp.analysis.bytes in_flight	数值类型	1.460×10^3
tcp.analysis.duplicate_ack_num	数值类型	1
tcp.dstport	布尔类型	2 090
tcp.flags.urg	布尔类型	0
tcp.len	数值类型	751
tcp.srcport	布尔类型	80
tcp.window size	数值类型	12 864
udp.dstport	布尔类型	47 666
udp.length	数值类型	97
udp.srcport	布尔类型	47 521

将布尔类型特征值字段转换为二进制值的格式后作为输入数据格式，其中，在输入深度学习网络模型时，将 TCP、UDP 等布尔类型特征值字段转换为二进制值的格式，作为输入数据格式。定义了一个 16 bit 的二进制列表，用于存储 TCP、UDP、HTTP 等端口号转换后的二进制值。将 frame、protocol 等文本类型的特征值字段，通过 BoW（bag of word，词汇假设）的方法进行格式转换后作为输入数据格式，将 Tcp.Len、Udp.Len 等数值类型特征值字段作为输入数据格式，转换后的特征值个数为 n'。为了便于处理数据量庞大的样本数据集并使之满足存储空间的可测量性要求，本文将散列法运用到 BoW 转换法中。为了提高梯度下降算法的效率，对各个特征值字段进行标准化规范：

$$z=\frac{x-\mu}{\delta} \tag{9}$$

其中，x 表示具体特征值，μ 和 δ 分别为特征值的期望和标准差，z 则为数据的标准化计算分数。

将特征转换后的 $m\times n'$ 的二维特征矩阵，用一系列连续的窗口尺寸为 T 的时间窗进行切割，并为每个时间戳设置标签值 y，标签值 y 为 0，表示该时间戳内数据分组为正常分组，标签值 y 为 1，表示该时间戳内数据分组为 DDoS 攻击分组。对切割后的特征进行维度重构，构建满足深度学习网络模型输入要求的 $(m-T)\times m\times n'$ 的三维矩阵。

3.3　模型检测阶段

模型检测阶段：首先构建包含输入层、卷积

层、递归神经网络、全连接层和输出层的深度学习网络模型，如图 5 所示。

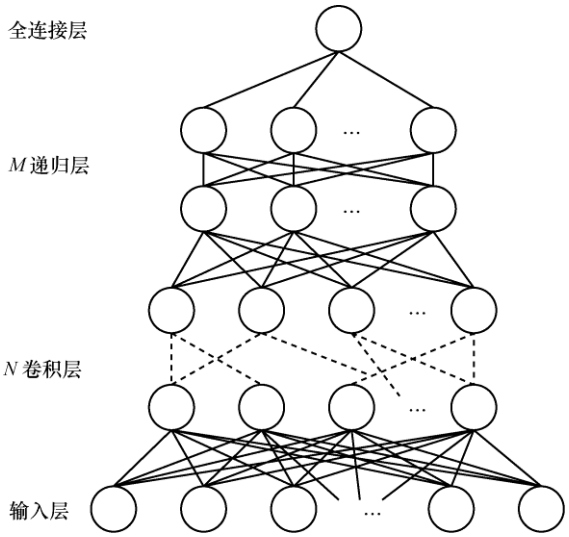


图 5　深度学习网络模型

其中，在输入层，采用批标准化加速神经元网络训练，同时为了获取本地信息和简化神经元网络，在后面的 N 层中采用堆栈卷积神经层，每层中有 128 个神经元，卷积神经网络的卷积核是 5，步幅为 1，且神经元的输出 f 作为非线性激活函数：

$$f(x)=\tanh x \tag{10}$$

卷积神经层之后，建立 M 层递归神经层。递归神经层能够帮助跟踪内存单元之前的时间戳。同时，为了解决梯度消失在深度 RNN 上的问题，本文尝试 LSTM 和 GRU（gated recurrent unit）。LSTM 旨在克服梯度消失在深度 RNN 上的问题和使用一个内存单元前时间戳。改进的 LSTM 通常包括 3 个门：输入、忘记和输出。GRU 是一个简单的变体，可以训练标准 LSTM 和更快，因为参数更少。在递归神经层上构建全连接层和设置它们的激活函数，其中在激活函数选取的过程中：若采用 Sigmoid 等函数，在反向传播求误差梯度时，求导涉及除法，计算量相对大，而采用 Relu 激活函数，整个过程的计算量会少很多；其次，对于深层网络，Sigmoid 函数反向传播时，很容易出现梯度消失的情况，会造成信息丢失。综合考虑，采用 Relu 作为激活函数 f：

$$f(x)=\max(0,x) \tag{11}$$

对于整个神经网络，在输出层将得到攻击分组的概率值。

本文基于不同的 LSTM 网络模型，构建多种深度学习网络模型，所构建的不同深度学习网络模型说明见表 2。

表 2　4 种不同类型的深度学习网络模型结构比较

模型名称	LSTM 模型	CNN/LSTM 模型	GRU 模型	3LSTM 模型
LSTM/GRU 层	4	4	4	6
神经元数目	64	64	64	64
激活函数	tanh	tanh	tanh	tanh
CNN 层	/	2	/	2
神经元数目	/	128	/	/
激活函数	/	Relu	/	/
全连接层	4	4	4	4
神经元数目	128,1	128,1	128,1	128,1
激活函数	Relu	Relu	Relu	Relu

由表 2 可知，根据基于的不同 LSTM 网络模型，所构建的深度学习网络模型也将有所区别。其中，对定义的 4 种不同类型的深度学习网络模型增加层数，综合考虑检测精度、误报率、计算时间等因素，最终确定实验中采用 LSTM 模型（8 层）、GRU 模型（8 层）、CNN/LSTM 模型（10 层）、3LSTM 模型（12 层）。

本文在向深度学习网络模型输入数据时，引入了时间戳的概念。通过这种方式，实现了一般 DDoS 攻击检测系统中基于单个数据报文输入的检测方式到基于窗口流量输入的 DDoS 攻击检测方式的转变。本文通过自定义改变时间戳，实现对网络环境中的 DDoS 攻击行为的检测，4 种模型采用不同时间戳，在检测精度上对比见表 3。

表 3　不同时间戳模型检测精度对比

窗口大小	步幅	LSTM 模型	CNN/LSTM 模型	GRU 模型	3LSTM 模型
500	50	89.50%	88.76%	90.56%	96.05%
1 000	50	87.73%	87.90%	88.24%	96.05%
5 000	50	98.85%	96.94%	90.32%	94.99%
500	10	98.41%	97.06%	96.46%	98.97%
1 000	10	98.73%	99.36%	98.69%	99.30%
500	5	99.48%	99.33%	99.65%	99.69%

4 种模型采用不同时间戳，在检测 F1 值上对比见表 4。

表 4　不同时间戳模型检测 F1 值对比

窗口大小	步幅	LSTM 模型	CNN/LSTM 模型	GRU 模型	3LSTM 模型
500	50	85.43%	84.86%	88.76%	90.13%
1 000	50	91.73%	89.90%	92.14%	95.35%
5 000	50	96.95%	94.64%	93.32%	92.79%
500	10	94.74%	96.86%	95.77%	97.81%
1 000	10	97.83%	97.86%	97.99%	98.53%
500	5	98.88%	98.23%	98.61%	99.02%

由于一个较大时间戳可以存储更长的攻击时间序列，并且可以反映更完整的攻击活动，因而时间戳的引入，对 DDoS 攻击行为的检测具有更强的说服力。通过改变检测窗口大小和步幅大小来改变时间戳，实验中时间戳选取 50 或 100，通过对比不同模型采用不同时间戳在检测精度和 F1 值上的表现，由表 3、表 4 可知，当窗口大小为 500、步幅为 5 时，即时间戳为 100，模型的检测精度和 F1 值最大，最终将时间戳默认设定为 100。

深度学习网络模型建立完成后，同时满足了实验要求的检测精度（本文在实验过程最低可达到 95%以上的检测精度），而且深度学习 DDoS 攻击检测模块加载满足要求的深度学习网络模型后，即可完成对 DDoS 网络攻击行为的检测。

4　实验与结果分析

4.1　实验环境

本文通过 ISCX 数据集训练检测模型，并通过实时的 DDoS 攻击对模型进行验证。其中，进行深度学习模型训练实验所基于的硬件环境为 2 个 NVIDIA K80 GPU 和 128 GB 内存，软件环境为 Ubuntu 14.04 操作系统和 Keras 深度学习框架，训练和检测所需的正常数据分组和 DDoS 攻击数据分组，通过 ISCX2012 数据集采集；实时的 DDoS 攻击检测所基于的硬件环境为 ATCA（advanced telecom computing architecture）机箱，其中 ATCA9700 40 GB 处理刀片，包含 2 个 10 核 CPU 和 8 条 DDR3-1886 内存条，ATCA3710 40 GB Fabric 交换刀片，包含一个 4 核 CPU 和一个 4 GB DDR3 SDRAM 内存条，包含 6 个 100 MB Base 接口通道和 12 个 10GbE SPF+光口通道。软件环境为 Centos6.5 操作系统和 Keras 深度学习框架。其中正常流量和 DDoS 攻击流量通过 Hping3 软件采集，Hping 是一个命令行下使用的 TCP/IP 数据分组组装/分析工具，不但能发送 ICMP 回应请求，还可以支持 TCP、UDP、ICMP 和 RAW-IP 等。实验过程中同时采用 Bandwidthd 监测网络流量状况，及时对 DDoS 攻击流量进行分析。

4.2　数据集训练检测模型

本次实验的输入样本数据时间戳为 100。为降低数据集的不可靠性，深度学习网络模型将进行 10 轮重复训练。在进行深度学习网络模型训练的过程中，每一轮训练过程，数据集的前 90%用于训练模型，剩余的 10%用于精度检测，即每轮先将训练集的前 90%输入深度网络模型进行训练，然后在训练结束后用剩余 10%的数据集作为本轮训练的验证数据集，重复 10 轮训练，从而提高模型的训练精度，并保存最终的深度学习网络模型。

本文使用 ISCX2012 数据集作为训练检测网络环境中 DDoS 攻击行为的深度学习网络模型的样本数据集。ISCX2012 记录的是 7 天时间的真实网络环境中的流量信息，其中包含合法的网络流量以及多种类型的恶意 DDoS 攻击流量。在这 7 天

的网络流量记录中，DDoS 攻击发生时间段为 2010 年 6 月 14—15 日，分别提取出这两天时间的网络流量并且将其保存到命名为 Sample_ Data_14 和 Sample_Data_15 两个文件当中。其中，Sample_Data_14 包含超过 960 万条数据报文信息，Sample_Data_15 则包含将近 3 500 万条数据报文信息。

ISCX2012 数据集列出了合法或各 DDoS 攻击类型的数据报文相关信息，包括数据报文类型名、抓取分组时间、源或目的 IP 地址、TCP/UDP 报文的源或目的端口号等字段信息。对 ISCX2012 数据集进行样本数据的分析统计，分别对 Sample_Data_14 和 Sample_Data_15 所记录的网络流量的统计分析，两个文件的主要攻击类型以及相应数据报文相关信息见表 5。

由表 5 可知，所列主要攻击类型及其数据报文出现的频率已知，首先，对网络流量中的每一个正常数据报文和攻击类型数据报文，根据已知攻击类型数据报文进行匹配并打上标签，每一条数据报文都将会被打上标签，标记该数据报文为攻击分组或非攻击分组。Sample_Data_14 和 Sample_Data_15 中大部分流量信息都是合法的数据分组流量，为了消除数据偏差，在输入深度网络模型进行训练时，本文采取的方案是每次在向深度学习网络模型输送数据时，将所有攻击类型数据报文与随机数目的合法数据报文混合，进行重新采样后再进行输入。

在深度学习网络模型训练过程，在采用相同数据集和特征值的情况下，通过比较基于不同 LSTM 网络模型构建的深度学习网络模型相关度量参数，从表 3 中选择相对最优的深度学习网络模型运用于深度学习 DDoS 攻击检测模块当中。比较过程中涉及的 5 种度量参数：分别为精度（accuracy）、准确率（precision）、召回率（recall）、F1 值（F1 score）和 AUC 值（AUC）。

为方便比较，将输入窗口尺寸进行统一设定。深度学习网络模型输入窗口越大，对网络环境中的 DDoS 攻击行为的检测越有效，故将窗口尺寸设定为 100，此时对 DDoS 攻击的检测精度最高，针对上述 4 种基于不同 LSTM 网络模型的深度学习网络模型，相应 5 种度量参数的比较结果见表 6。

由表 6 所列训练精度的结果可以看出，在 Sample_Data_14 数据集中，LSTM 深度学习网络模型在 5 种性能度量参数中表现最好，对于 Sample_Data_15 数据集，3LSTM 深度学习网络模型的 5 种性能度量参数表现最好。结合上述训练结果以及现实网络环境中 DDoS 攻击的大规模流量状况可以看出，3LSTM 深度学习网络模型针对大规模数据流量的检测，具有更好的效果。

基于 LSTM 网络模型的 3LSTM 深度学习网络模型训练，分别对 Sample_Data_14 和 Sample_Data_15 样本数据集进行训练结果验证，验证结果如图 6、图 7 所示，其中，图 6 记录了

表 5　分布式拒绝服务攻击流量信息说明

Sample_Data_14		Sample_Data_15	
DDoS 攻击类型	出现频率	DDoS 攻击类型	出现频率
IMAP 攻击	32	/	/
泛洪	50	安全 Web 服务	1
网络控制报文协议	55	DNS 解析	1
误用	77	邮件传输协议	1
安全 Web 服务	92	未知 TCP	5
未知 TCP	896	IRC 攻击	212
HTTP 访问	2 165	HTTP 访问	37 158

表 6　4 种不同深度学习网络模型性能比较

数据集	模型名称	精度	准确率	召回率	F1 值	AUC 值
Sample_Data_14 数据集	LSTM 模型	97.886%	98.118%	98.001%	98.102%	99.305%
	CNN/LSTM 模型	95.942%	97.397%	93.909%	95.729%	98.378%
	GRU 模型	96.742%	98.419%	95.098%	97.091%	99.012%
	3LSTM 模型	96.657%	98.011%	95.463%	97.919%	99.159%
Sample_Data_15 数据集	LSTM 模型	96.939%	97.954%	97.892%	97.989%	99.324%
	CNN/LSTM 模型	96.501%	96.305%	97.182%	97.521%	99.184%
	GRU 模型	98.097%	98.213%	98.004%	98.126%	99.335%
	3LSTM 模型	98.501%	98.422%	98.674%	98.496%	99.602%

Sample_Data_14 样本数据集在 10 轮重复训练过程中，精度值随训练次数变化而发生变化的关系；同理，图 7 则记录了 Sample_Data_15 样本数据集在 10 轮重复训练过程中，精度值随训练次数变化而发生变化的关系。

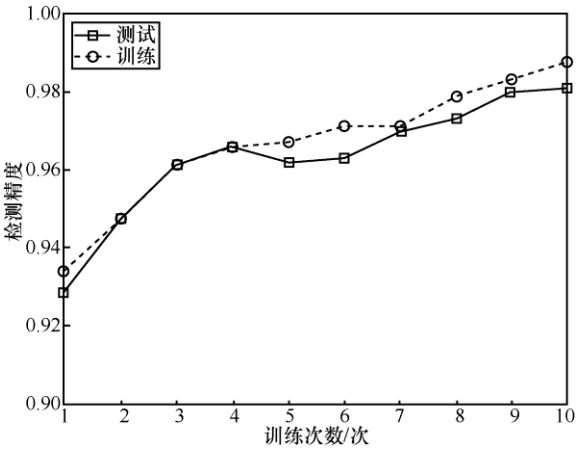


图 6　Sample_Data_14 样本数据集训练精度

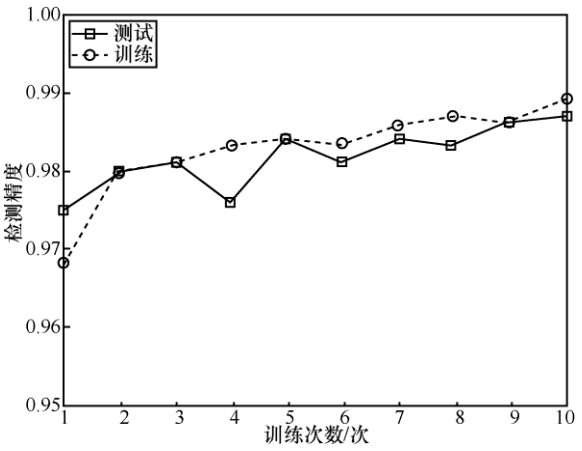


图 7　Sample_Data_15 样本数据集训练精度

从图 6、图 7 模型训练及验证结果可以看出，对样本数据集进行 10 次重复训练并检验，随着训练次数的增加，训练精度值越高，对数据集进行验证的精度也越高。其中，在第 10 次重复训练时，模型训练和检测对应的精度值达到最高，Sample_Data_14 数据集的最高精度趋于 98%，而 Sample_Data_15 数据集的最高精高达 99%，由此可见，增加重复训练的次数，换言之，增大训练样本数据集的规模，将会使得深度学习网络训练和验证的精度显著提升。

本文比较了深度学习和参考文献[12]中所提到的随机森林方法检测 DDoS 攻击，调整窗口大小，对两种模型检测精度进行测试，结果如图 8 所示。

由图 8 可知，随机森林模型虽然能够区分正

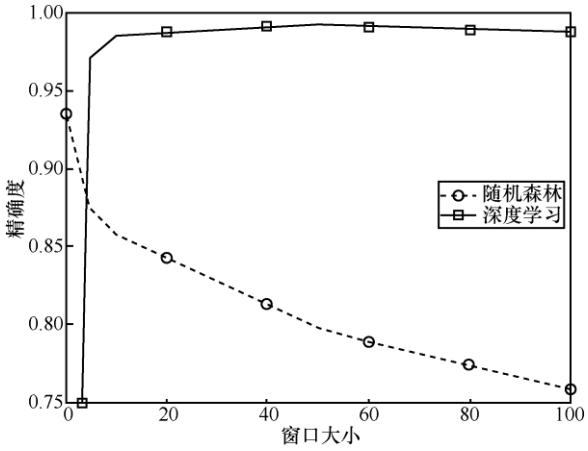


图 8　深度学习与随机森林检测精度

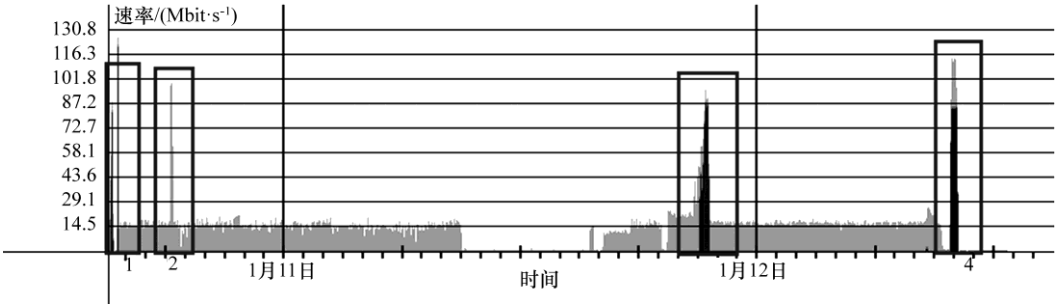


图 9　eth0 网卡流量信息

常流量和攻击流量，但是这种方式只适用于针对单个分组的检测，对于多重数据分组，检测准确率较低。随着检测窗口的增大，深度学习模型的检测精度表现更好，随机森林模型出现了内存溢出的现象，检测精度急剧下降。

4.3　实时 DDoS 攻击检测验证

基于训练好的模型检测实时 DDoS 攻击，同时深度学习网络模型将分别进行 10 轮测试，并采用统计的方法测试模型检测的准确率。实验过程中采用 Bandwidthd 软件监测 2017 年 1 月 10 日 15:00—1 月 12 日 10:15 内 ATCA 机箱中 eth0 网卡的流量信息，如图 9 所示。

图 9 中在 1、2、3、4 处所标注的时间段内，采用 Hping3 发送了不同类型的 DDoS 攻击，显然，当 DDoS 攻击发生时，流量出现爆炸性增长，最大达到 135 Mbit/s。

SYN 泛洪（SYN flood）攻击，是一种攻击者通过向被攻击目标发送虚假报文以欺骗被攻击服务器的攻击方式。实际网络中，如果服务器的 TCP/IP 协议栈容量有限，不够强大，最终通常导致堆栈溢出崩溃。为了验证模型实时检测 DDoS 攻击的能力，采用 Hping3 软件发送 SYN 泛洪攻击，图 9 中 2 处表明网络中 TCP 流量变化情况，最大速率达到 97 Mbit/s，实验过程中采用发送的数据分组总数量为 1 000 000 个，其中 SYN 泛洪攻击分组为 950 000 个，正常分组为 50 000 个，正常数据分组的发送速率恒定为 500 个/s，进行 10 次检验，并提高攻击分组的发送速率，模型的检测精度如图 10 所示。

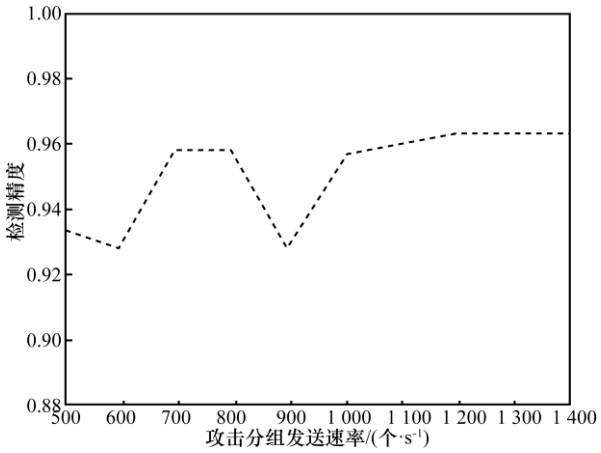


图 10　SYN 泛洪攻击检测精度

Smurf 型攻击是攻击者向被攻击端发送一种源 IP 地址设置为被攻击端的 IP 地址，目的 IP 地址为广播地址的畸形 ICMP 数据报文的 DDoS 攻击方式，被攻击端收到 ICMP 数据报文后，将回复 ICMP 应答报文。ICMP 泛洪（ICMP flood）攻击是一种流量型 DDoS 攻击方式，它利用规模庞大的数据流量，造成被攻击目标超负荷运行，从而导致被攻击目标无法正常完成服务业务功能；UDP 泛洪（UDP flood）攻击利用这一特性，向被攻击目标发送大量的 UDP 数据报文，最终导致被攻击目标所在网络因充斥大量无效的数据流量而耗尽带宽资源。为了验证深度学习模型对不同类型的 DDoS 攻击的检测精度，使用 Hping3 分别发送 SYN 泛洪、ICMP 泛洪、UDP 泛洪、LAND 攻击（LAND attack）、XMAS 攻击（XMAS tree）、Smurf 攻击，图 9 中 1、2 处表示出现不同类型的 DDoS 攻击时，网络中 TCP、UDP、ICMP 类型的流量变

化，其中 UDP 类型流量最大速率达到 85 Mbit/s，ICMP 类型流量最大速率达到 120 Mbit/s，TCP 类型流量最大速率达到 135 Mbit/s。实验过程中采用发送的数据分组总数量为 1 000 000 个，其中 DDoS 攻击分组为 950 000 个，正常分组为 50 000 个，正常分组的发送速率恒定为 500 个/s，每种类型的攻击进行 10 次实验，并提高攻击分组的发送速率，对于不同类型的 DDoS 攻击，模型的检测精度如图 11 所示。

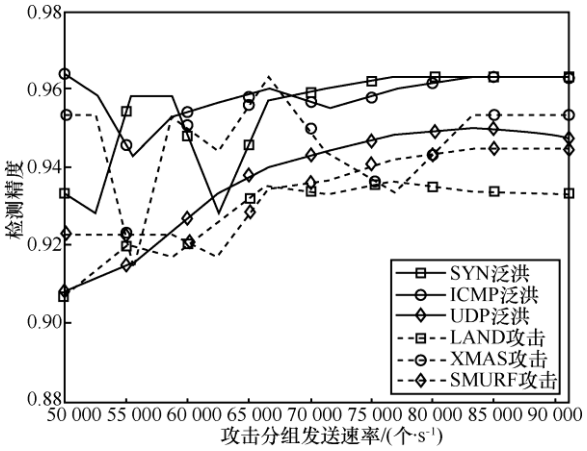


图 11　不同种类的 DDoS 攻击检测精度

图 11 中验证了深度学习模型对 SYN 泛洪、ICMP 泛洪、UDP 泛洪、LAND 攻击、XMAS 攻击、SMURF 攻击 6 种不同类型的 DDoS 攻击的检测精度，其中，模型对 SYN 泛洪、ICMP 泛洪、XMAS 攻击的检测进度最大达到 96.3%，对 UDP 泛洪检测精度最大达到 95%，对 LAND 攻击检测精度最大达到 93.6%，对 SMURF 检测精度最大达到 94.5%。

为了验证深度学习模型对组合类型的 DDoS 攻击的检测精度，使用 Hping3 分别发送 SYN 泛洪+ICMP 泛洪、SYN 泛洪+UDP 泛洪、UDP 泛洪+ICMP 泛洪攻击，并提高攻击分组的发送速率，图 9 中 3、4 处表示出现组合类型的 DDoS 攻击时，网络中 TCP、UDP、ICMP 类型的流量变化，最大速率达到 114 Mbit/s。实验过程中采用发送的数据分组总数量为 1 000 000 个，其中 DDoS 攻击分组为 950 000 个，正常分组为 50 000 个，且正常数据分组的发送速率恒定为 500 个/s，每种类型的攻击进行 10 次实验，并提高攻击分组的发送速率，对于不同组合 DDoS 攻击类型，模型的检测精度如图 12 所示。

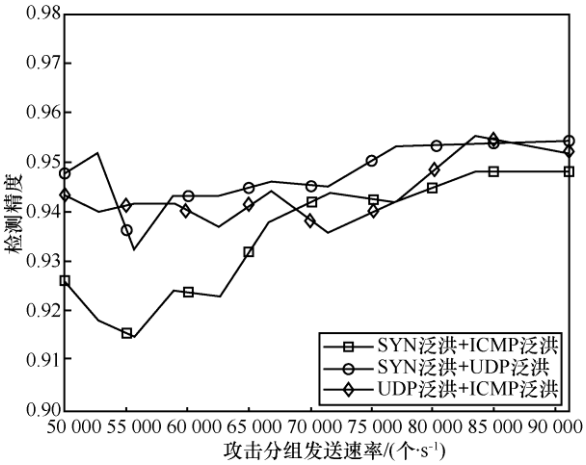


图 12　组合型 DDoS 攻击类型检测精度

图 12 中验证了深度学习模型对 SYN 泛洪+ICMP 泛洪、SYN 泛洪+UDP 泛洪、UDP 泛洪+ICMP 泛洪攻击 3 种不同组合类型的 DDoS 攻击的检测精度。其中，模型对 SYN 泛洪+ICMP 泛洪攻击的检测进度最大达到 94.8%，对 SYN 泛洪+UDP 泛洪检测精度最大达到 95.4%，对 UDP 泛洪+ICMP 泛洪检测精度最大达到 93.6%，对 SMURF 检测精度最大达到 95.5%。

深度学习模型是基于窗口大小检测数据分组是否为 DDoS 攻击分组的，试验中设置窗口大小为 100，为了验证正常数据分组在模型检测窗口中的占空比时验证深度学习模型的检测精度，实验过程中采用发送的数据分组总数量为 1 000 000 个，使用 Hping3 发送 SYN 泛洪攻击，其中 DDoS 攻击分组为 900 000 个，且攻击分组的发送速率为 50 000 个/s，正常分组为 100 000 个，且提高正常数据分组的发送速率，即提高正常分组在模型检测窗口中的占空比，模型的检测精度如图 13 所示。

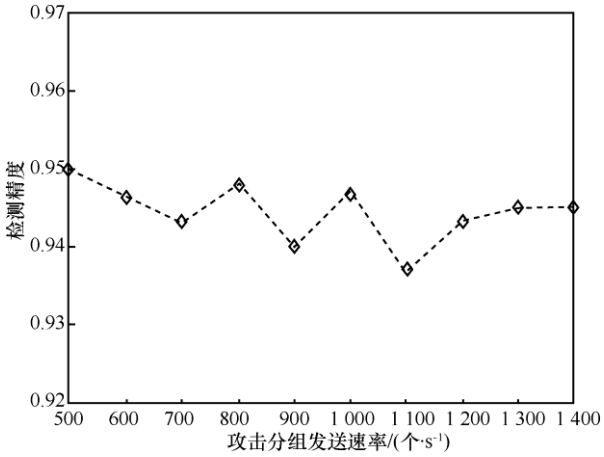


图 13 SYN 泛洪攻击检测精度

如图 13 所示，提高正常数据分组的在深度学习检测模型窗口中的占空比，深度学习模型对 SYN 泛洪攻击检测精度的最大值为 95%，随着正常数据分组在检测窗口中的占空比增大，模型的检测精度总体呈现下降的趋势，并且低于同种类型的 DDoS 攻击的检测精度，表明正常分组在检测窗口中的占空比对模型的检测精度有影响。

5 结束语

本文所提出的基于深度学习的 DDoS 攻击方法，从最终的实验结果来看，模型训练阶段使用 ISCX 数据集对 DDoS 攻击的最终的验证精度高达 98%甚至 99%以上，且深度学习模型和随机森林模型相比在多数据分组检测方面更具有优势。在实时 DDoS 攻击对模型验证，最高精度达到 96.3%。由此可以看出，基于深度学习的 DDoS 攻击检测方案具有非常好的检验效果，结合实验过程中的软硬件环境，可以看出基于深度学习的 DDoS 检测方案具有高检测精度、对软硬件设备依赖小、深度学习网络模型易于更新等优点，弥补了现有 DDoS 攻击检测方案的不足。简而言之，基于深度学习的 DDoS 攻击检测方案的优点可以总结为：在提高 DDoS 攻击检测精度的同时也降低了系统对软硬件环境的依赖程度，同时简化了检测系统实时更新和 DDoS 攻击检测策略的升级难度。

深度学习网络模型在针对大规模 DDoS 攻击网络流量的训练和检测时，实时 DDoS 攻击检测精度将有所下降，模型的检测精度与 DDoS 攻击分组发送速率、DDoS 攻击类型和正常分组在检测窗口中的占空比有关系。今后的研究中将改进模型具有更高的准确度和对检测环境的适应性，DDoS 攻击流量越大，深度网络模型训练的精度也将越高，对 DDoS 攻击检测的准确性也将大大提高。

参考文献：

[1] PENG T, LECKIE C, RAMAMOHANARAO K. Survey of network-based defense mechanisms countering the DoS and DDoS problems[J]. ACM Computing Surveys, 2007, 39(1): 3.

[2] MIRKOVIC J, MARTIN J, REIHER P. A taxonomy of DDoS attacks and DDoS defense mechanisms[J]. ACM Sigcomm Computer Communication Review, 2001, 34(2): 39-53.

[3] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[J]. Advances in Neural Information Processing Systems, 2012, 25(2): 1097-1105.

[4] GRAVES A, SCHMIDHUBER J. Framewise phoneme classification with bidirectional LSTM and other neural network architectures[J]. Neural Networks, 2005, 18(5): 602-610.

[5] DEY R, FATHI M. Gate-variants of gated recurrent unit (GRU) neural networks[EB/OL]. (2017-01-28)[2017-03-14]. http://xueshu.baidu.com/s?wd=Gate-Variants+of+Gated+Recurrent+Unit+%28GRU%29+Neural+Networks&rsv_bp=0&tn=SE_baidux-ueshu_c1gjeupa&rsv_spt=3&ie=utf-8&f=8&rsv_sug2=1&sc_f_para=sc_tasktype%3D%7BfirstSimpleSearch%7D&rsv_n=2.

[6] UNB ISCX intrusion detection evaluation DataSet[EB/OL]. (2014-05-08)[2017-03-14]. http://www.unb.ca/research/iscx/dataset/iscx-IDS-dataset.html.

[7] BHUYAN M H, BHATTACHARYYA D K, KALITA J K. Information metrics for low-rate DDoS attack detection: a comparative evaluation[C]//International Conference on Contemporary Computing, August 7-9 , 2014, Noida, India. New Jersey: IEEE Press, 2014: 80-84.

[8] CHEN Y, KAI H. Collaborative change detection of DDoS attacks

on community and ISP networks[C]//International Symposium on Collaborative Technologies and Systems, May 14-17, 2006, Las Vegas, NV, USA. New Jersey: IEEE Press, 2006: 401-410.

[9] YUAN J, MILLS K. Monitoring the macroscopic effect of DDoS flooding attacks[J]. IEEE Transactions on Dependable & Secure Computing, 2005, 2(4): 324-335.

[10] BHUYAN M H, BHATTACHARYYA D K, KALITA J K. An empirical evaluation of information metrics for low-rate and high-rate DDoS attack detection[J]. Pattern Recognition Letters, 2015, 51(C): 1-7.

[11] CHEN C L. A new detection method for distributed denial-of-service attack traffic based on statistical test[J]. Journal of Computer Science, 2009, 15(2): 488-504.

[12] SINGH K J, DE T. An approach of DDoS attack detection using classifiers[M]. Berlin: Springer, 2015: 429-437.

[13] LECUN Y, BOTTOU L, BENGIO Y, et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11): 2278-2324.

[14] KAMIJO K, TANIGAWA T. Stock price pattern recognition-a recurrent neural network approach[C]//International Joint Conference on Neural Networks, June 17-21, 1990, San Diego, CA, USA. New Jersey: IEEE Press, 1990: 215-221.

[15] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. Neural Computation, 1997, 9(8): 1735-1780.

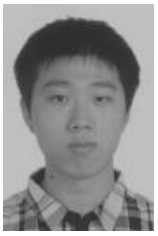
[16] 邬江兴. 拟态计算与拟态安全防御的原意和愿景[J]. 电信科学, 2014, 30(7): 2-7.
WU J X. Meaning and vision of mimic computing and mimic security defense[J]. Telecommunications Science, 2014, 30(7): 2-7.

[17] 柳毅, 洪俊斌. 基于网络爬虫与页面代码行为的 XSS 漏洞动态检测方法[J]. 电信科学, 2016, 32(3): 87-91.
LIU Y, HONG J B. A dynamic detection method based on Web crawler and page code behavior for XSS vulnerability[J]. Telecommunications Science, 2016, 32(3): 87-91.

[作者简介]

李传煌（1980−），男，博士，浙江工商大学信息与电气工程学院副教授，2016 年佛罗里达大学访问学者，主要研究方向为软件定义网络、深度学习、开放可编程网络、系统性能预测和分析模型，发表 EI/SCI 检索论文 40 余篇，申请专利 15 项。

孙正君（1993−），男，浙江工商大学信息与电气工程学院硕士生，主要研究方向为软件定义网络、深度学习。



袁小雍（1990−），男，美国佛罗里达大学博士生，主要研究方向为网络安全、深度学习、云计算和分布式系统。

李晓林（1976−），男，美国佛罗里达大学副教授，大规模智能系统实验室（Large-Scale Intelligent Systems Laboratory，Li lab）的创始人，美国 NSF I/UCRC CBL 中心（Center for Big Learning，CBL）主任，主要研究方向为云计算、大数据、深度学习、SDN、健康及精准医药学、CPS/IoT 等，获得美国国家科学基金（NSF）、国家卫生研究院（NIH）、国土安全部（DHS）等的大力资助，发表期刊和会议论文 100 余篇，出版专著 4 部，申请 4 项美国发明专利（3 项被授权）。

龚梁（1992−），男，浙江工商大学信息与电气工程学院硕士生，主要研究方向为网络安全、深度学习、软件定义网络。

王伟明（1964−），男，博士，浙江工商大学信息与电子工程学院教授，主要研究方向为新一代网络架构、开放可编程网络，特别是 IETF ForCES、SDN 及可重构网络等方面的协议、模型和算法。