2 LSTM 网络模型设计

工控网络尤其是电力系统的网络使用场景单一, 在网络安全运行的情况下, 流量数据表现平稳, 并且具有周期性, 一旦网络发生异常, 流量就会产生较大的波动, 具体表现为流量各个维度的数值相较于历史数据发生突变, 不再符合周期性的特点. 基于工控网络流量平稳和周期性强的特点, 本文提出使用 LSTM 网络模型对工控网络的流量数据进行时序预测[11−14], 在网络正常运行的情况下, 可以认为模型的预测值为正常值, 当某一时刻的实际值偏离预测值较大, 即认为网络出现异常.

2.1 异常流量检测流程

检测流程分为两个阶段, 第一个阶段是解析流量数据包构建有效特征, 第二个阶段是 LSTM 网络模型的离线训练过程和在线检测过程. 第一阶段在电力 SCADA 系统中通过镜像端口的方式采集通信流量, 对采集到的数据包进行深度包解析, 针对 104 规约数据包, 除了要解析常规的源地址、目的地址、源端口、目的端口、标志位和连接时间等基本信息, 还要解析长度为 6 个字节的 APCI, 里面包含了控制电路的操作信息. 对解析到的字段进行整合并重新构造出模型输入需要的特征, 除了采集流量构造出的特征, 构建时序模型还需要加入滞后历史特征, 即模型需要用多长时间的输入去预测下一时刻的数值. 第二阶段在模型离线训练完成后, 即可部署线上环境, 对电力通信网络进行异常流量预测并实时报警.

2.2 特征构造

如图 1 所示, 第一阶段包含从流量中提取特征, 并构造有效特征两部分, 流量分为两部分, 第一部分是离线采集的流量, 主要用于模型训练, 第二部分是源源不断的在线流量. 从流量中解析出同 TCP/IP 协议一样的字段和 104 规约中的 APCI 字段. 解析字段分为 3 类, 第一类为 9 个内部属性, 这些属性是从网络数据包的头部中提取得到, 例如连接的持续时间 (duration), 连接的协议类型 (protocol_type), 包含 http、ftp、smtp 和 telnet 等 70 种网络服务类型 (service) 等; 第二类为内容属性, 这些属性是从网络数据包的内容区域中提取, 例如从应用规约数据单元中提取的信息体、数据单元标识和 104 规约报文变长帧中的 APCI; 第三类为派生属性, 这些属性的计算考虑了之前的连接, 细分为时间流量属性和机器流量属性, 时间流量属性考虑过去 2 秒内发生的连接, 例如到同一目标 IP 地址的点击总和 (count), 到同一目标端口号的连接总和 (srv_count) 等字段. 对以上 38 个字段构建时间统计特征, 构建方法分为计数 (count)、占比 (percent) 和均值 (average), 共构造出 12 个有效特征. 例如“相同主机”特征, 检查过去 2 秒内与当前连接具有相同目标主机的连接, 统计连接的数量, 再计算与当前连接具有相同服务的连接百分比、不同服务的百分比、SYN (泛洪) 的百分比以及 REJ(拒绝连接) 的百分比. 对于诸如报文内部属性中的网络服务类型等的非数值型特征, 需要对其进行独热编码 (one-hot) 转换为数值型特征用于模型的输入.

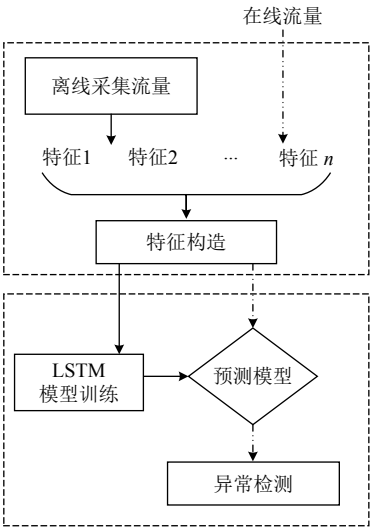

图 1 异常流量检测流程

2.3 模型训练

图 2 中的第二阶段为模型训练部分, 分为离线训练和在线检测两部分, 离线训练把构造好的特征输入到初始化的 LSTM 网络中进行模型训练, 训练好的预测模型输入以相同方式构造的特征样本进行异常流量检测, 如果模型检测到网络中的流量异常则发出警报.

3 LSTM 网络结构和参数更新

3.1 网络整体结构

LSTM 网络结构由 RNN 加入门控机制改进得到, RNN[15] 能够很好地处理不固定长度并且有序的输入序列. RNN 前向传播过程如图 3 所示, 网络参数权重的更新不仅仅依赖每一时刻t样本输入x_t对参数w的调整, 而且依赖t时刻之前计算并保存的隐含状态h_{t-1}对参

数的调整. 与传统的 RNN 相比, LSTM[16] 本质上还是基于t时刻的输入x_t和$t-1$时刻的隐状态h_{t-1}来计算t时刻的输出y_t和t时刻的隐状态h_t. 但是由于门控机制的加入, LSTM 网络更适合处理长依赖问题, 更加容易学习到工控网络周期性的规律, 并且容易识别由多个数据包共同作用引起的攻击类型.

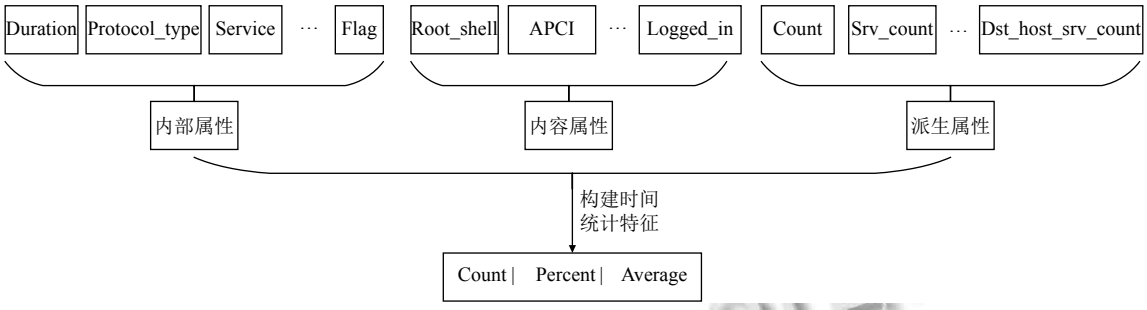

图 2 特征构造流程

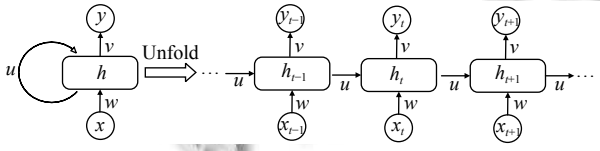

图 3 RNN 前向传播过程

本文提取数据包字段并构造了 12 个有效的时间统计特征, 网络模型在t时刻的结构为一个简单的前馈神经网络, 整体的网络结构如图 4 所示, 有 N 个前馈神经网络组成, 不同时刻的前馈神经网络通过隐藏层神经元传递依赖关系. 每层的前馈神经网络分为 3 层, 分别是包含 12 个神经元节点的输入层 (Input Layer), 含有 64 个神经元节点的隐藏层 (Hidden Layer), 含有 12 个神经元的输出层 (Output Layer), 在训练过程中, 前 N 个时刻的流量数据包用于预测 N+1 时刻的流量统计值, 即前 N 个时刻为样本特征, 第 N+1 个时刻为样本标签.

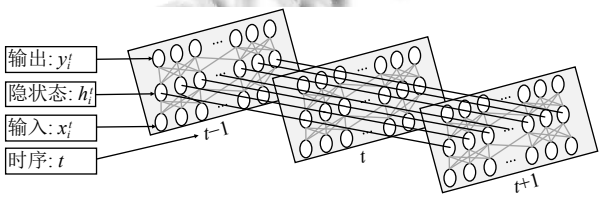

图 4 LSTM 神经网络时间展开图

3.2 参数更新过程

LSTM 网络相比 RNN 增加了存储单元用来存储长期记忆, 增加了输入门用来记忆 t 时刻的输入信息, 新来一个样本, 并不会完全学习记忆其中的特征, 而是自动学习除其中有多少有用信息可以用于 N+1 时刻的预测. 遗忘门用来选择性的忘记过去的某些信息, 起控制内部信息的作用. 输出门起控制输出信息的作用, 3 个门控单元的加入让 LSTM 网络在用梯度下降算法更新参数时不易于陷入梯度消失的问题, 3 个门的逻辑结构如图 5 所示.

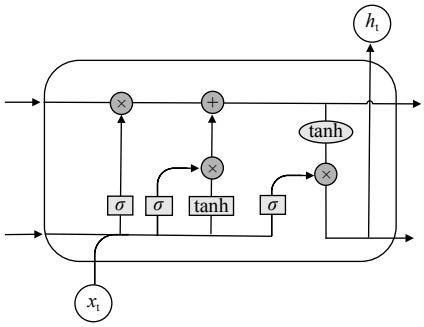

图 5 LSTM 网络门控机制

输入门输入 15 分钟以内 180 个样本的时间统计值, 15 分钟为滞后历史特征数值, 在训练过程中是一个超参数, 经过多组训练实验得到最优滞后历史特征, 对滞后历史特征数值的选择如表 1 所示, 可以发现, 当滞后历史特征为 15 分钟时, 模型在验证集上的损失最低, 表面用前 15 分钟的流量去预测下一时刻流量的时间统计特征最准确, 本文以 5 s 为最小单位, 在预测流量时, 预测下一时刻 (5 s 内) 的流量统计特性. 网络内部输入门的计算过程为

$$i_t = \sigma(w_i \cdot [h_{t-1}, x_t] + b_i) \quad (1)$$

其中, σ为 Sigmoid 激活函数. 遗忘门、输出门和输入

© 中国科学院软件研究所 http://www.c-s-a.org.cn

门计算方式一样, 细胞状态C_t用于长期记忆, 更新过程为:

$$C_t = f_t * C_{t-1} + i_t * C_t \tag{2}$$

隐状态h_t用于短期记忆, 更新过程为:

$$h_t = o_t * \tanh(C_t) \tag{3}$$

其中, o_t为输出门的输出, tanh 为双曲正切激活函数.

表 1 滞后历史特征和隐藏层神经元节点数训练情况

滞后历史特征 (分钟)	隐藏层神经元节点数	Train_loss (训练误差)	Validation_loss (验证误差)
10	600	0.84	0.75
12	800	0.91	0.87
15	960	0.98	0.97
18	1100	0.98	0.95
20	1200	0.98	0.93

本文的隐藏层神经节点有 960 个, 网络输入的是一个三维向量 [640,180,12], 第一维 batch_size 的含义是一次性将 640 个样本序列, 输入到网络中进行训练, 使用梯度下降的方法完成一次误差反向传播和参数更新, 第二维 time_step 的含义是用前 180 个样本去预测下一时刻的流量值, 第三维 input_size 是单个样本的维度. 网络的输出是一个二维向量 [640×180, 12], 第一维代表输出 (预测) 的时刻流量值, 第二维代表单个样本的维度. 网络的输出为数值型数据, 所以损失函数采用均方误差损失函数[17], 定义为

$$MSE(y,y') = \frac{\sum_{i=1}^{n}(y_i - y_i')^2}{n} \tag{4}$$

其中, y为样本标签, y'为模型预测值, 输出值为 1×12 的向量, 图 6 为网络输出层的数据流图, 隐藏输出为输出层的输入, 经过 reshape 后和输出层权重进行点积运算, 加上偏置后得到 115 200 个样本的预测值.

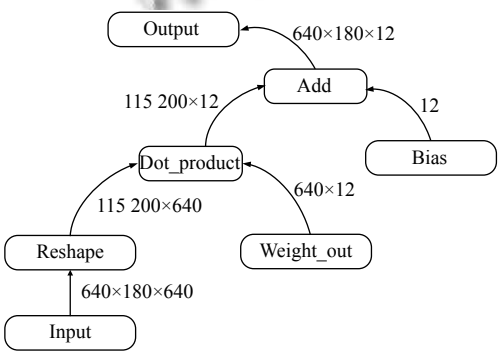

图 6 网络输出层数据流图

4 实验设计与结果分析

本文训练模型所用的数据采集自东北电力公司某子网, 利用 C++库函数 libpcap 对数据包进行捕获和深度解析, 对捕获到的数据包在时间维度上进行整合, 对时间间隔为 5 s (慢速攻击标准) 内的流量报文构造统计特征生成一个样本, 数据集的大小为 4.26 GB, 将数据集按照采集时间分为训练集和验证集, 训练集和数据集的比例为 70%:30%, 由于时间序列的原因, 划分数据集不能随机打乱, 而是按照采集流量的时间线, 把前 70% 的数据包划分为训练集, 后 30% 的数据包划分为验证集, 供模型训练和验证其有效性[18].

4.1 模型训练过程

模型训练过程中, 网络参数可以由训练得到, 滞后历史特征和隐藏层神经节点个数两个超参数通过网格调参的方式选取最优组合, 通过多轮训练的结果, 如表 1 所示, 可以发现, 最优组合为滞后历史特征的值为 15 分钟, 隐藏层的神经元节点数的值为 960 个. 模型每更新 500 次参数后计算一次训练误差和验证误差, 在迭代到第 19 轮时, validation_loss (验证误差) 达到最小, 在最优组合的超参数下, 模型在验证集上的准确率可以达到 97%, 图 7 为模型的 validation_loss 下降过程, 从图 7 中可以看出, 从第 19 次迭代后, 模型的 validation_loss 不再下降.

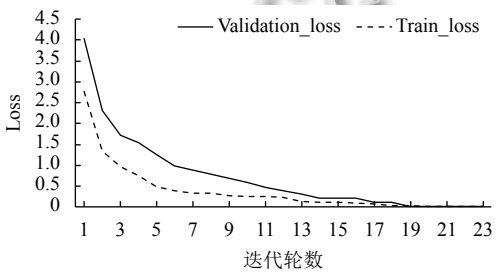

图 7 训练和验证损失

4.2 实验结果对比

表 2 列举出了多种主流算法对工控网络异常流量的识别率、误报率和识别效率, 可以发现, 本文算法在识别率和识别效率均优于半监督的 K-means 算法、单类支持向量机 (One Class Support Vector Machine, OCSVM) 和 BP 神经网络, 相比卷积神经网络方法, 本文算法误报率稍高, 但是识别效率却快了几倍. 总体而言, 本文算法结合工控网络周期性强和流量报文具有时序的特点, 使用 LSTM 模型取得了较好的效果.

© 中国科学院软件研究所 http://www.c-s-a.org.cn

表 2 各类算法测试结果对比

指标	K-means 聚类	OCSVN	BP神经网络	卷积神经网络	本文算法
识别率(%)	87.01	88.62	92.37	97.88	97.67
误报率(%)	12.99	10.67	8.05	3.39	9.37
识别用时(s)	21	18	24	26	6

5 总结

本文以工控网络中的电力系统网络为研究对象, 使用 LSTM 算法识别工控网络流量异常, 结合工控网络场景相较单一和周期性的特点, 采集流量后对解析的数据包字段解析重构时间统计特征, 采用时序预测的方式识别流量异常, 通过实验可以发现, 能有效识别出异于正常情况的网络波动, 由于提前预测出正常流量的特征值, 算法在异常流量的识别效率上优于传统识别方法, 有利于技术人员尽早发现异常做出相应的安全防护措施, 提高工控网络在入侵检测方面的安全性. 本文提出的时序预测模型虽然在识别准确率和识别效率相较其它算法有所提升, 但是时序预测模型要求流量数据具有周期性这一特点, 并且模型的最终效果非常依赖训练前期的特征构造, 目前特征构造中使用计数、百分比和均值统计指标, 后续为了进一步提高模型的识别率, 降低模型的误报率, 会在特征构造中加入其它的统计指标.

参考文献

1 周民军. 工控网络现状与安全分析. 现代工业经济和信息化, 2017, (15): 61–62.

2 吴震生. 基于工控协议防火墙的脆弱性研究. 自动化与仪表, 2019, 34(8): 105–108.

3 崔秀帅. 智能变电站报文安全及其实时性研究 [硕士学位论文]. 哈尔滨: 哈尔滨工业大学, 2016.

4 倪震. 电力工控网络安全风险分析与预测关键技术研究 [博士学位论文]. 南京: 南京理工大学, 2018.

5 安攀峰. 基于 SVM 的工业控制网络入侵检测方法应用研究. 电子世界, 2017, (1): 154, 156.

6 刘进, 卢安文, 陈家佳. 基于模糊聚类算法的工控网络控制故障检测方法研究. 计算机测量与控制, 2015, 23(3): 681–684. [doi: 10.3969/j.issn.1671-4598.2015.03.002]

7 祝士祥. 基于深度学习的工控网络异常检测研究和实现 [硕士学位论文]. 北京: 北京邮电大学, 2017.

8 张艳升, 李喜旺, 李丹, 等. 基于卷积神经网络的工控网络异常流量检测. 计算机应用, 2019, 39(5): 1512–1517. [doi: 10.11772/j.issn.1001-9081.2018091928]

9 张望妮, 牛瑞. 利用 IEC104 远动规约解决实际问题. 电子技术与软件工程, 2018, (21): 159–160.

10 江泽鑫. IEC60870-5-104 规约安全性分析及攻击实验. 信息技术与网络安全, 2018, 37(10): 1–4, 14.

11 张照宇. 时序事件分析与研究 [硕士学位论文]. 西安: 西安电子科技大学, 2019.

12 曾豪. 基于 LSTM 的环境污染时间序列预测模型的研究 [硕士学位论文]. 武汉: 华中科技大学, 2019.

13 王伟. 基于深度学习的网络流量分类及异常检测方法研究 [博士学位论文]. 合肥: 中国科学技术大学, 2018.

14 卜国卿. 网络流量异常检测技术研究与实现 [硕士学位论文]. 成都: 电子科技大学, 2018.

15 Liu RR. Short-term traffic flow prediction based on deep circulation neural network. Proceedings of the 2nd International Seminar on Artificial Intelligence, Networking and Information Technology(ANIT 2018). Shanghai, China. 2018. 580–583.

16 Navares R, Aznarte JL. Predicting air quality with deep learning LSTM: Towards comprehensive models. Ecological Informatics, 2020, 55: 101019. [doi: 10.1016/j.ecoinf.2019.101019]

17 董巧玲. 不同误差影响模型下总体最小二乘法在多元线性回归中的应用研究 [硕士学位论文]. 太原: 太原理工大学, 2016.

18 陈凯. 深度学习模型的高效训练算法研究 [博士学位论文]. 合肥: 中国科学技术大学, 2016.

© 中国科学院软件研究所 http://www.c-s-a.org.cn