基于多源异构数据融合的网络安全态势评估体系

常利伟， 田晓雄， 张宇青， 钱宇华， 胡治国

引用本文:
常利伟, 田晓雄, 张宇青, 等. 基于多源异构数据融合的网络安全态势评估体系[J]. 智能系统学报, 2021, 16(1): 38–47.
CHANG Liwei, TIAN Xiaoxiong, ZHANG Yuqing, et al. Network security situation assessment architecture based on multi-source heterogeneous data fusion[J]. *CAAI Transactions on Intelligent Systems*, 2021, 16(1): 38–47.

在线阅读 View online: https://dx.doi.org/10.11992/tis.202006053

您可能感兴趣的其他文章

基于注意力机制的显著性目标检测方法
Salient object detection method based on the attention mechanism
智能系统学报. 2020, 15(5): 956–963 https://dx.doi.org/10.11992/tis.201903001

基于Hadoop的大规模网络安全实体识别方法
Large-scale network security entity recognition method based on Hadoop
智能系统学报. 2019, 14(5): 1017–1025 https://dx.doi.org/10.11992/tis.201809024

多特征融合的lncRNA识别与其功能预测
LncRNA recognition by fusing multiple features and its function prediction
智能系统学报. 2018, 13(6): 928–934 https://dx.doi.org/10.11992/tis.201806008

城市轨道交通线网数据中心与评估决策平台
A platform for a data center and decision making in urban rail transit
智能系统学报. 2018, 13(3): 458–468 https://dx.doi.org/10.11992/tis.201612005

改进D-S证据理论在电动汽车锂电池故障诊断中的应用
Application of improved D-S evidence theory in fault diagnosis of lithium batteries in electric vehicles
智能系统学报. 2017, 12(4): 526–537 https://dx.doi.org/10.11992/tis.201605001

基于脑连接网络的阿尔茨海默病临床变量值预测
Prediction of clinical variables in Alzheimer's diseaseusing brain connective networks
智能系统学报. 2017, 12(3): 355–361 https://dx.doi.org/10.11992/tis.201607020

微信公众平台

关注微信公众号，获取更多资讯信息

DOI: 10.11992/tis.202006053

基于多源异构数据融合的网络安全态势评估体系

常利伟[1,2], 田晓雄[1], 张宇青[1], 钱宇华[2], 胡治国[2]

(1. 山西财经大学 信息学院, 山西 太原 030006; 2. 山西大学 大数据科学与产业研究院, 山西 太原 030006)

摘 要: 针对基于单点网络数据很难准确地检测网络恶意活动且无法有效地分析网络状况的问题, 本文通过引入多源异构数据融合策略, 借鉴层次化网络分析思想, 构建出包含流量探测模块、属性提炼模块、决策引擎模块、多源融合模块、态势评估模块等五大模块的网络安全态势评估体系。评估体系以 BP 神经网络为决策引擎分析各数据源的数据, 使用指数加权 D-S 证据理论融合各决策引擎的输出结果, 并基于层次化网络威胁评估方法评估网络威胁状况。实验结果表明: 不同探测器探测到的数据对于识别不同类型攻击的优势不同; 多源融合技术进一步将识别攻击类型的准确率提升到 88.7%; 层次化网络威胁评估方法能够有效地评估网络威胁状况。

关键词: 网络安全; 网络安全态势评估; 数据融合; 层次化分析方法; 网络攻击; 威胁量化; 检测评估

中图分类号: TP393 **文献标志码**: A **文章编号**: 1673−4785(2021)01−0038−10

中文引用格式: 常利伟, 田晓雄, 张宇青, 等. 基于多源异构数据融合的网络安全态势评估体系 [J]. 智能系统学报, 2021, 16(1): 38–47.

英文引用格式: CHANG Liwei, TIAN Xiaoxiong, ZHANG Yuqing, et al. Network security situation assessment architecture based on multi-source heterogeneous data fusion[J]. CAAI transactions on intelligent systems, 2021, 16(1): 38–47.

Network security situation assessment architecture based on multi-source heterogeneous data fusion

CHANG Liwei[1,2], TIAN Xiaoxiong[1], ZHANG Yuqing[1], QIAN Yuhua[2], HU Zhiguo[2]

(1. College of Information, Shanxi University of Finance and Economics, Taiyuan 030006, China; 2. Institute of Big Data Science and Industry, Shanxi University, Taiyuan 030006, China)

Abstract: Because it is difficult to detect malicious network activity precisely and analyze the network situation effectively based only on the single point network data, in this paper, we propose a network security situation assessment architecture consisting of five modules: a traffic detection module, attribute extraction module, decision engine module, multi-source fusion module, and situation assessment module based on the strategy of multi-source heterogeneous data fusion and the idea of hierarchical network security assessment. In this assessment architecture, a BP neural network is used as the decision engine to analyze the multi-source heterogeneous data, the exponential weighting D-S evidence theory is used to merge the output of multiple decision engines, and the threat status of the network is exhibited by referring to the hierarchical network security threat assessment method. The experimental results demonstrate that first, the data from different detectors have different advantages for identifying different types of attacks; second, the multi-source fusion technology can further improve the accuracy of identifying attacks, which is up to 88.7%; and third, the hierarchical network analysis method can exactly exhibit the threat status of network effectivity.

Keywords: network security; network security situation assessment; data fusion; hierarchical analysis method; network attacks; threat quantification; detection and evaluation

收稿日期: 2020−06−30.

基金项目: 山西省自然科学基金项目 (201801D221159); 山西省高等学校科技创新项目 (2019L0470); 山西省重点研发项目 (201903D421003).

通信作者: 常利伟. E-mail: changliwei002@163.com.

没有网络安全就没有国家安全, 网络安全已成为信息时代国家安全的基石。然而随着网络规模的日益扩大以及网络恶意行为的复杂化与智能

化, 传统以入侵检测系统为核心的单点防御体系暴露出越来越多的弊端。单点检测方式本质上是通过单个节点的信息做出判断, 会形成“安全信息孤岛”, 无法从整个网络层面做出准确的决策, 因此面对复杂的网络恶意活动时会产生大量的误报、漏报。

安全态势评估通过技术手段从时间和空间维度来感知并获取安全相关元素, 综合分析安全信息以准确判断安全状况。随后国内外许多研究人员将态势评估的思想及方法应用到网络安全领域, 设计和实现了许多高效网络安全态势评估系统。

1995 年 Endsley[1] 将态势感知概括为态势觉察、态势理解、态势投射 3 个过程。1999 年 Bass[2] 首次提出网络安全态势感知, 本质上是融合多源入侵检测系统的结果, 识别网络中的攻击活动, 评估网络运行状况。2002 年陈继军[3] 提出权系数理论原则, 用于确定各传感器的系数。2005 年诸葛建伟等[4] 引入 D-S 证据理论, 构建决策引擎判断网络状况。2006 年陈秀真等[5] 提出了层次化安全威胁评估模型, 通过网络流量信息及入侵检测报警信息, 对网络运行状况进行评估。2008 年马琳茹等[6] 通过指数加权规则, 将不同的传感器赋予不同的信任以改进 D-S 证据理论。2009 年韦勇等[7] 从节点脆弱性和攻击威胁等角度, 基于 D-S 证据理论算法, 构建网络安全态势评估模型。

2016 年刘效武等[8] 构建了一个融合−感知−决策−控制的态势认知融合感控模型, 对网络状况进行评估。2018 年 Wang Huan 等[9] 使用混淆矩阵最大特征值对应的特征向量确定主机态势的参数, 对网络运行状况进行评估。2018 年龚俭等[10] 认为网络安全态势感知 (network security situation awareness, NSSA) 是认知网络系统安全状态的过程, 包含原始数据测量、语义提取、融合处理、异常识别、态势获取等内容。2018 年 Zhao Dongmei 等[11] 通过粗糙集属性简约算法提取核心属性, 并使用粒子群优化算法优化径向基神经网络识别网络攻击。2018 年陈维鹏等[12] 将网络态势等级进行划分, 通过模拟退火算法优化 BP(back propagation) 神经网络参数, 确定网络空间态势感知等级。2019 年贾焰等[13] 对网络安全态势感知概念、网络安全态势关键技术等方面的研究现状进行深入剖析。2019 年 Xi Rongrong 等[14] 从威胁、脆弱性和稳定性 3 个维度评估网络的安全状况, 并在决策层面将结果进行融合来衡量整个网络的安全状况。2020 年 Zheng Weifa 等[15] 使用 D-S 证据理论融合主机防火墙数据、web 防火墙数据和入侵检测数据, 对网络安全性进行评估。

通过对以上学术成果的综合分析, 本文构成出基于多源异构数据融合的网络安全态势评估体系, 主要工作如下:

1) 提出了包含流量探测模块、属性提炼模块、决策引擎模块、多源融合模块、态势评估模块等 5 大模块的网络安全态势评估体系。

2) 以入侵检测系统 (Snort) 报警规则为标尺, 提炼网络威胁等级划分原则。

3) 使用层次化网络安全威胁评估方法, 依次评估服务层、主机层、网络层态势。

1　网络安全态势评估体系

本文提出的基于多源异构数据融合的网络安全态势评估体系包含流量探测模块、属性提炼模块、决策引擎模块、多源融合模块和态势评估模块等 5 大模块。如图 1 所示, 5 个模块构建时吻合信息安全管理中安全基础、识别认定、检测评估、安全防护 4 个维度, 充分考虑可用性、可控性、保密性等信息安全要求。如图 2 所示: 1) 流量探测模块是在网络中部署多个探测器, 以全面地获取网络信息; 2) 属性提炼模块是基于恶意活动特征, 准确地构造有助于提高识别攻击类型的核心属性; 3) 决策引擎模块是利用网络数据, 科学地训练由核心属性到攻击类型的模型; 4) 多源融合模块是巧妙地融合决策引擎的输出结果, 有效提升识别攻击类型的性能; 5) 态势评估模块是基于融合结果, 直观地展示网络运行状态。

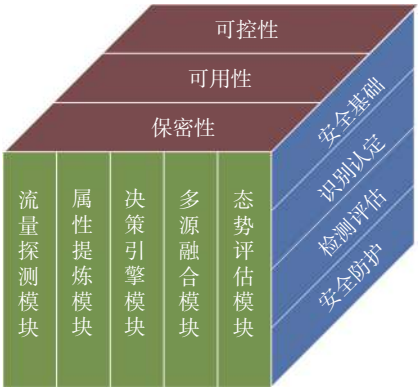

图 1　网络安全态势评估体系

Fig. 1　The architecture of network security situation assessment

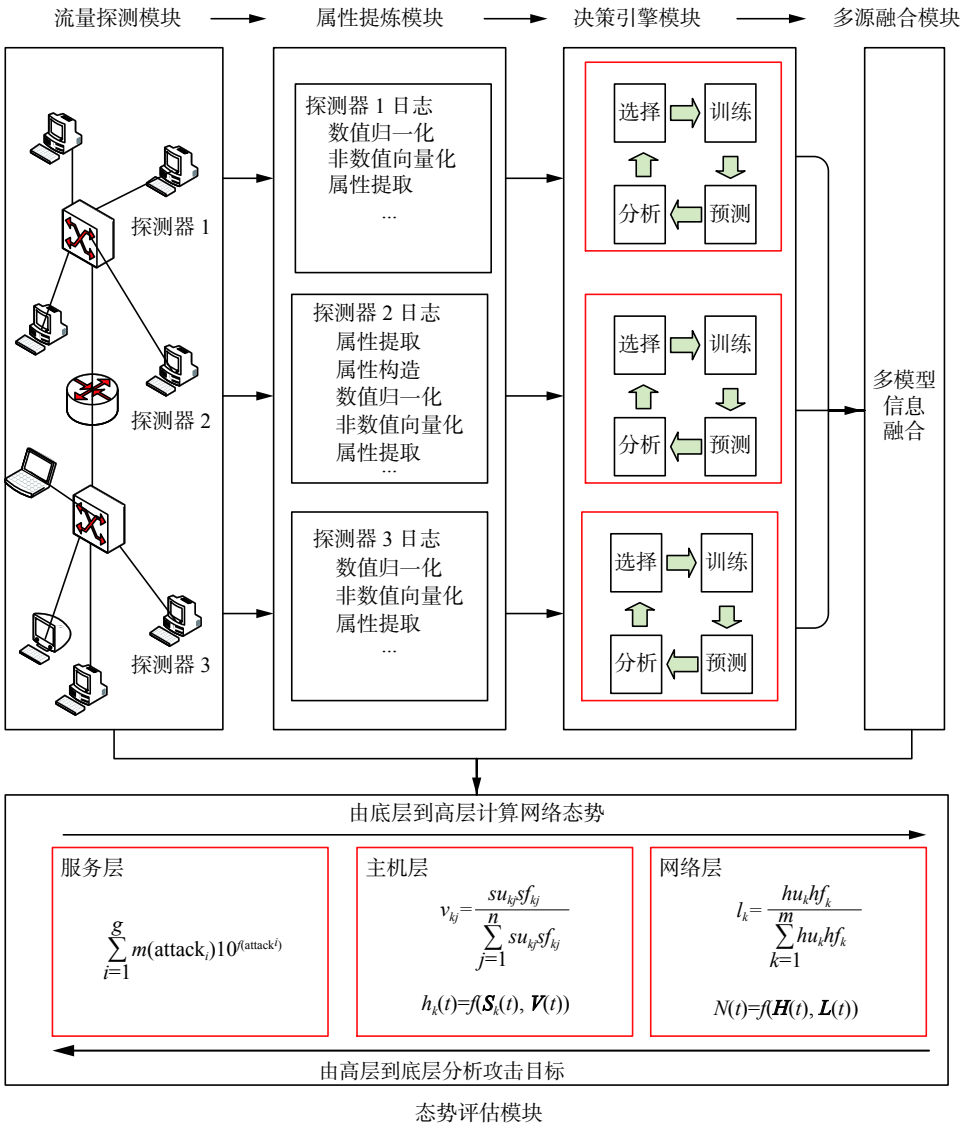

图 2　模块组件关系

Fig. 2　The relationship of modules

1.1　流量探测模块

信息在网络中以流量包为单元，在 2 台设备之间进行传递，流量包由一台主机发出，途径路由器、交换机、防火墙、网卡等设备，到达另一台主机。流量探测模块将在以上设备中部署多个探测器，尽可能全面地获取网络数据。

常用的探测器分以下 2 类：1) 网络流量探测器，这类探测器通常部署在局域网入口路由器上，可以全面地、实时地获取所有流入、流出局域网的网络流量，并发送至数据处理中心；2) 入侵检测系统，入侵检测系统实时检测网络流量信息，并与规则库中的报警规则进行匹配，一旦发现异常流量，将发出对应的警告信息。

1.2　属性提炼模块

属性提炼模块根据网络恶意活动的特征，准确地构造有助于提高识别攻击类型的核心属性。

Netflow、Snort、Suricata 探测器获得的基础属性如表 1~3 所示。1) 将其中数值型数据进行归一化处理，避免数值变化较大的属性覆盖数值较小的属性，使得数值变化较小的属性失去作用；2)将非数值型属性编码成向量，使得计算机能够处理。

由于其侧重点各异，不同探测器将获取到不同属性的网络数据。因此本文设计不同的统计算法，以高效地提炼不同探测器的属性：1) 研究发现如果直接使用 Netflow 和 Suricata 中的地址属性、端口属性、应用服务属性和时间属性进行攻击识别，效果不明显，因此借助统计算法融合以上 4 类属性生成网络连接属性[16-19](如表 4)，其他属性保留；2) Snort 和 Suricata 为入侵检测系统，其所产生的报警信息是其核心要素，因此针对性地设计统计算法提炼报警数量与报警类别属性。

表 1　Netflow 基础属性
Table 1　Basic features of Netflow

序号	属性名称	描述
1	Stime	开始时间
2	Dur	持续时间
3	Ltime	结束时间
4	Protocol	传输层协议
5	Sip	源ip地址
6	Dip	目的ip地址
7	Sport	源端口号
8	Dport	目的端口号
9	Pkt	包数
10	Byt	包大小
11	Bps	比特数/s
12	Pps	包数/s
13	Bpp	平均包大小

表 2　Snort 基础属性
Table 2　Basic features of Snort

序号	属性名称	描述
1	Sip	源ip地址
2	Dip	目的ip地址
3	Sport	源端口号
4	Dport	目的端口号
5	Ip_len	Ip包长度
6	Attack_num	报警数量
7	Attack_type	报警类型

表 3　Suricata 基础属性
Table 3　Basic features of Suricata

序号	属性名称	描述
1	Stime	开始时间
2	Ltime	结束时间
3	Sip	源ip地址
4	Dip	目的ip地址
5	Sport	源端口号
6	Dport	目的端口号
7	Sever	应用层服务
8	Pkt	包数
9	Byt	包大小
10	Attack_num	报警数量
11	Attack_type	报警类型

表 4　网络连接属性
Table 4　Statistical features of network traffic

序号	属性名称	描述
1	is_sm_ips_prots	Sip与Dip相同并且Sport与Dport相同为1, 否则为0
2	ct_dst_ltm	每100条记录中, Ltime、Dip相同的数量
3	ct_src_ltm	每100条记录中, Ltime、Sip相同的数量
4	ct_src_dport_ltm	每100条记录中, Ltime、Sip、Dport相同的数量
5	ct_dst_sport_ltm	每100条记录中, Ltime、Dip、Sport相同的数量
6	ct_dst_src_ltm	每100条记录中, Ltime、Sip、Dip相同的数量
7	ct_srv_src	每100条记录中, Sip、Serve相同的数量
8	ct_srv_dst	每100条记录中, Dip、Serve相同的数量

1.3　决策引擎模块

决策引擎模块的功能是通过有效训练, 准确地学习出从各探测器核心属性到攻击类型的模型。神经网络是一个优秀的分类模型, 学者们常用此模型作为决策引擎[11, 20], 本文采用 BP 神经网络作为决策引擎。

BP 神经网络由输入层、隐藏层、输出层构成。Robert Hecht Nielson[21] 证明, 只包含一个隐藏层的网络可以逼近闭合区间内的任一连续函数, 因此 BP 神经网络能够实现由属性到攻击类型的映射。

含有 1 层隐藏层的 BP 神经网络训练过程如下, 输入层有 m 个神经元, 隐藏层有 h 个神经元, 输出层有 n 个神经元, 激活函数为 $f(x)$, 隐藏层神经元输出为 $\boldsymbol{D}=(d_1,d_2,\cdots,d_h)^{\mathrm{T}}$, 输入层与隐藏层之间的权值为 $\boldsymbol{W}\{w_{ji}|1\leqslant i\leqslant m,1\leqslant j\leqslant h\}$, 隐藏层与输出层之间的权值为 $\boldsymbol{V}\{v_{kj}|1\leqslant k\leqslant n,1\leqslant j\leqslant h\}$, 神经网络预测值为$\widehat{\boldsymbol{Y}}=(\widehat{y}_1,\widehat{y}_2,\cdots,\widehat{y}_n)^{\mathrm{T}}$, 真实值为 $\boldsymbol{Y}=(y_1,y_2,\cdots,y_n)^{\mathrm{T}}$, 均方误差为 $\boldsymbol{E}$, 如图 3 所示。

隐藏层第 j 个节点的输出值为

$$d_j=f\left(\sum_{i=1}^{m}x_iw_{ji}\right),1\leqslant j\leqslant h \tag{1}$$

输出层第 k 个节点的输出值为

$$\widehat{y}_k=f\left(\sum_{j=1}^{h}d_jv_{kj}\right),1\leqslant k\leqslant n \tag{2}$$

均方误差为

$$\boldsymbol{E}=\frac{1}{2}\sum_{k=1}^{n}(\widehat{y}_k-y_k)^2 \tag{3}$$

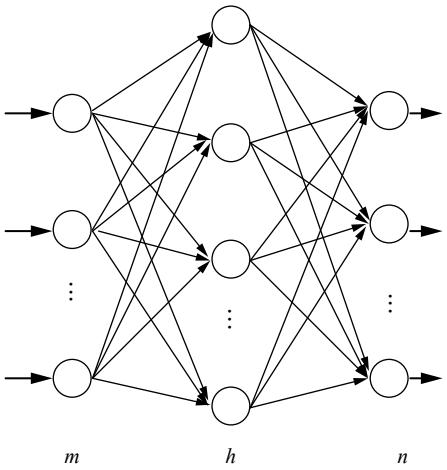

图 3　BP 神经网络

Fig. 3　BP netural network

输出层及隐藏层各神经元的权值调整为

$$\Delta v_{kj} = -\eta\frac{\partial \boldsymbol{E}}{\partial v_{kj}} = -\eta\frac{\partial \boldsymbol{E}}{\partial \widehat{y_k}}\frac{\partial \widehat{y_k}}{\partial v_{kj}} = -\eta(\widehat{y_k} - y_k)f'\left(\sum_{j=1}^{h} d_j v_{kj}\right)d_j \tag{4}$$

$$\Delta w_{ji} = -\eta\frac{\partial \boldsymbol{E}}{\partial w_{ji}} = -\eta\left(\sum_{k=1}^{n}\frac{\partial \boldsymbol{E}}{\partial \widehat{y_k}}\frac{\partial \widehat{y_k}}{\partial d_j}\right)\frac{\partial d_j}{\partial w_{ji}} = -\eta\left(\sum_{k=1}^{n}(\widehat{y_k} - y_k)f'\left(\sum_{j=1}^{h} d_j v_{kj}\right)v_{kj}\right)f'\left(\sum_{i=1}^{m} x_i w_{ji}\right)x_i \tag{5}$$

1.4　多源融合模块

多源融合模块将有机地融合所有决策引擎的输出结果，进一步提高识别攻击的性能。融合算法将各个决策引擎的输出结果作为融合因子，将相同安全事件组合处理得到融合结果。本文使用指数加权 D-S 证据理论融合决策引擎输出结果，并利用粒子群优化算法确定指数权重 (particle swarm optimization-dempster shafer, PSO-DS)。

指数加权 D-S 证据理论为

$$m(A_i) = \frac{m_1(A_i)^{w_{1i}} m_2(A_i)^{w_{2i}} \cdots m_g(A_i)^{w_{gi}}}{K} \tag{6}$$

$$K = \sum_{i=1}^{n} m_1(A_i)^{w_{1i}} m_2(A_i)^{w_{2i}} \cdots m_g(A_i)^{w_{gi}} \tag{7}$$

式中：g 为数据源个数；n 为攻击类型个数；$m_1(A_i)$ 为第一个证据源认为是第 i 类攻击的概率；w_{1i} 为第一个证据源认为是第 i 类攻击的指数权值；$m(A_i)$ 为融合后第 i 类攻击的概率。

使用粒子群优化算法搜索指数权值：

$$v_{td} = c_3 v_{td} + c_1 r_1(p_{td} - x_{td}) + c_2 r_2(g_{td} - x_{td}) \tag{8}$$

$$x_{(t+1)d} = x_{td} + v_{td} \tag{9}$$

式中：d 为优化空间维度；c_1、c_2 分别为将粒子推向局部最优和全局最优的权重；c_3 为惯性权重；r_1, r_2 为随机数。y_i 为真实数据中第 i 种攻击的概率，粒子群优化目标为

$$\arg\min\left(\sum_{i=1}^{n}(m(A_i) - y_i)^2\right) \tag{10}$$

1.5　态势评估模块

态势评估模块的功能是基于多源融合模块的输出结果，有效地评估网络态势。可分为攻击威胁量化和态势评估 2 个内容。

1.5.1　攻击威胁量化

量化影响态势变化的关键因子 (如：攻击威胁因子)，是科学评估网络运行状态的基础。其中攻击威胁因子量化是目前研究的一个难点，本文使用权系数理论量化攻击威胁。经过划分威胁等级和威胁等级量化 2 个步骤得到攻击威胁因子值。

1) 划分威胁等级

Snort 作为一个常用的入侵检测系统，受到用户的广泛认可。根据危害大小，Snort 能够对攻击威胁定性分析，表 5 列出了攻击威胁等级的核心内容。深入剖析威胁等级划分机理，本文提出了攻击威胁等级划分原则。

表 5　网络攻击威胁程度

Table 5　Severity of network attack

报警类型	描述	威胁等级
attempted-admin	成功获取管理员控制权限	高
attempted-user	成功获取普通用户控制权限	高
policy-violation	潜在的侵犯企业隐私活动	高
shellcode-detect	检测到可执行代码	高
attempted-dos	企图发动拒绝服务攻击	中
denial-of-service	检测到拒绝服务攻击	中
successful-dos	成功发动拒绝服务攻击	中
successful-recon-limited	信息泄露	中
icmp-event	通常发生的ICMP事件	低
misc-activity	杂乱的网络活动	低
network-scan	检测到网络扫描	低
not-suspicious	没有可疑流量	低

①威胁等级划分机理

如表 5 所示，攻击严重程度为高的攻击基本上是获取计算机控制权限、木马、执行代码等恶意活动，此类攻击会威胁主机系统安全；攻击严重程度为中的攻击以获取系统内部信息和消耗网络带宽为目的，这类攻击不会对主机系统造成破坏；严重程度为低的攻击以网络扫描、获取网络信息为目的，这类攻击对网络造成较轻影响。

②威胁等级划分原则

通过对 Snort 划分攻击威胁等级机理的深入研究, 本文提出了攻击威胁等级划分原则, 如表 6 所示。

表 6　威胁等级划分原则

Table 6　Classification principles of attack severity

类别	描述	威胁等级
计算机权限	获取计算机控制权限	高
隐私信息	获取系统内部信息	中
带宽消耗	消耗网络带宽	中
网络扫描	扫描获取网络信息	低

2) 威胁等级量化

本文将权系数理论[3] 与攻击威胁等级划分原则有机结合以量化威胁等级。权系数分布函数如图 4 所示, 横轴是排队等级 (威胁等级), 纵轴是对应的权系数 (威胁值)。

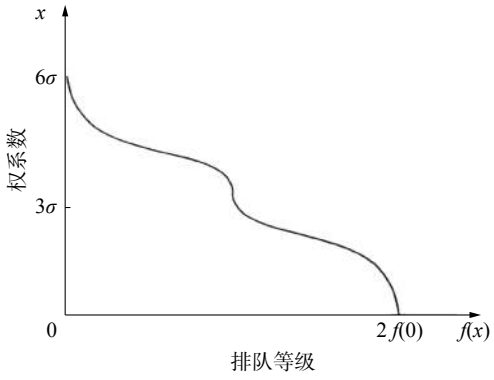

图 4　权系数函数分布

Fig. 4　Distribution of the weight function

为简化量化过程, 将权系数函数离散化处理得到权系数公式, n 为攻击等级个数, i 为排列顺序, 权系数只与决策目标数和攻击严重等级有关。仅需定义威胁等级, 使用权系数公式就能够计算出攻击威胁因子值。在一定程度上减轻了主观依赖问题。

权系数为

$$W_i=\begin{cases}\dfrac{1}{2}+\dfrac{\sqrt{-2\ln\dfrac{2i}{n}}}{6}, & 1\leqslant i<\dfrac{n}{2}\\[2ex] \dfrac{1}{2}, & i=\dfrac{n}{2}\\[2ex] \dfrac{1}{2}-\dfrac{\sqrt{-2\ln\left(2-\dfrac{2i}{n}\right)}}{6}, & \dfrac{n}{2}<i<n\end{cases}\tag{11}$$

1.5.2　态势评估

本文采用层次化网络安全威胁评估模型, 将网络自底向上分为服务层、主机层和网络层。按照层次关系, 以攻击对网络造成的危害程度为核心评估安全态势, 如图 5 所示。

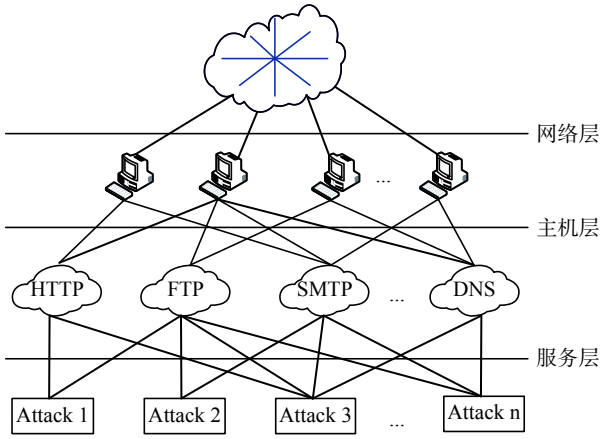

图 5　层次化网络安全威胁态势评估

Fig. 5　Hierarchical threat assessment model for computer networks

1) 服务层态势

攻击威胁因子和攻击概率等态势因子会影响服务态势, 服务态势随着网络运行状况实时发生变化, 在时间窗口 Δt 内第 k 台主机上第 j 个服务的态势为

$$s_{kj}(t)=\sum_{i=1}^{g}m(\text{attack}_i)\,10^{f(\text{attack}_i)}\tag{12}$$

式中: g 为该时间窗口内的攻击总数; $m(\text{attack}_i)$ 为第 i 个攻击对应的攻击概率; $f(\text{attack}_i)$ 为第 i 个攻击对应的攻击威胁因子值。

2) 主机层态势

主机的态势情况由主机上运行的服务态势及其在主机上的重要性决定, 在时间窗口 Δt 内, 第 k 台主机的态势值为

$$h_k(t)=f(\boldsymbol{S}_k(t),\boldsymbol{V}(t))=\boldsymbol{S}_k(t)\,\boldsymbol{V}_k(t)\tag{13}$$

$\boldsymbol{S}_k(t)=(s_{k1}(t),s_{k2}(t),\cdots,s_{kn}(t))$ 为此时间段内运行在该主机上服务态势的向量, 其中 n 为运行的服务数, $\boldsymbol{V}_k(t)=(v_{k1},v_{k2},\cdots,v_{kn})$ 为该主机上运行服务的权值向量。服务权值计算方法为

$$v_{kj}=\frac{su_{kj}sf_{kj}}{\sum_{j=1}^{n}su_{kj}sf_{kj}}\tag{14}$$

$\boldsymbol{Su}_k(t)=(su_{k1},su_{k2},\cdots,su_{kn})$ 表示该主机上使用应用层服务的用户数量的向量, $\boldsymbol{S}f_k(t)=(sf_{k1},sf_{k2},\cdots,sf_{kn})$ 为该主机上各个服务使用频率向量。

3) 网络层态势

网络层态势是由网络中每台主机的态势及每台主机的权值决定的, 整个网络的态势为

$$\boldsymbol{N}(t)=f(\boldsymbol{H}(t),\boldsymbol{L}(t))=\boldsymbol{H}(t)\boldsymbol{L}(t) \tag{15}$$

$\boldsymbol{H}(t)=(h_1(t),h_2(t),\cdots,h_m(t))$ 为网络中主机的态势向量，$\boldsymbol{L}(t)=(l_1,l_2,\cdots,l_m)$ 为主机的权值，网络中主机数量为 m。

$$l_k=\frac{hu_k hf_k}{\sum_{k=1}^{m} hu_k hf_k} \tag{16}$$

式中：m 为该网络中的主机数量，$\boldsymbol{Hu}(t)=(hu_1,hu_2,\cdots,hu_m)$ 为网络中各个主机用户数量的向量；$\boldsymbol{H}f(t)=(hf_1,hf_2,\cdots,hf_m)$ 为网络中各个主机使用频率的向量。

通过态势评估策略，将网络的安全状况及其演化状况准确地计算出来。如若发生威胁，管理员可以根据网络态势变化曲线及时地掌握网络状况，根据层次图由上到下依次分析网络层态势−主机层态势−服务层态势，追根溯源，准确快速的定位问题根源，从而采取有效地措施，保护网络安全。

2　实验分析

如图 6 所示，根据实验需求，部署 Netflow，以获取网络中流量的信息 (如端口号、流量包大小、发包速度等)，并部署 Snort、Suricata 等入侵检测工具，当发现异常流量则发出警告信息。实验中 BP 神经网络包含 2 层隐藏层，每层 32 个神经元。粒子群优化算法群体规模为 100，在 [0,1] 内搜索确定 D-S 指数权重。

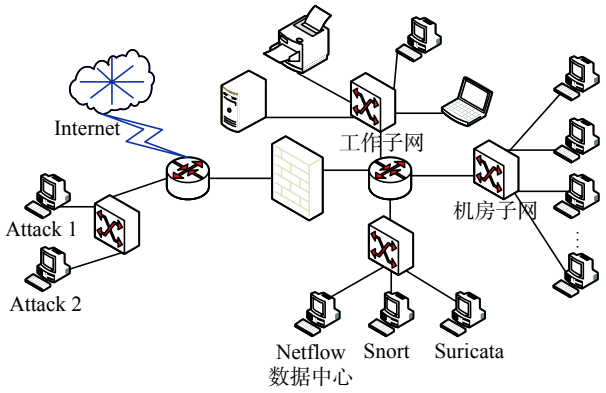

图 6　网络拓扑

Fig. 6　network topology

流量数据使用澳大利亚新南威尔逊大学安全实验室的 UNSW-NB15 流量包的 5% 用作本实验的数据，此流量包包含 2 天的流量包和攻击信息，混合了正常网络流量和综合性攻击流量，更符合现代真实的流量场景，而且网络规模更大，主机数量更多，攻击类型更加丰富，如表 7所示。

表 7　攻击类型及威胁因子

Table 7　Attack types and risk factors

序号	攻击名称	威胁程度	威胁因子
1	Analysis	中	0.611
2	Backdoor	高	0.726
3	Dos	中	0.611
4	Exploit	高	0.726
5	Fuzzers	高	0.726
6	Generic	中	0.611
7	Reconnaissance	低	0.389
8	Shellcode	高	0.726
9	Worm	高	0.726
10	Normal	—	0

2.1　决策引擎结果与融合性能分析

图 7~12 展示了决策引擎和融合算法的实验结果。分析可知来自不同探测器数据对于识别各类攻击各有优势，Netflow 数据源对于识别 Analysis、Normal 准确率较高；Snort 数据源对于区分 Backdoor、Worm 攻击效果很好；Suricata 可以很好地识别 Dos。融合算法集成了各数据源识别攻击的优势，提高了决策效果。

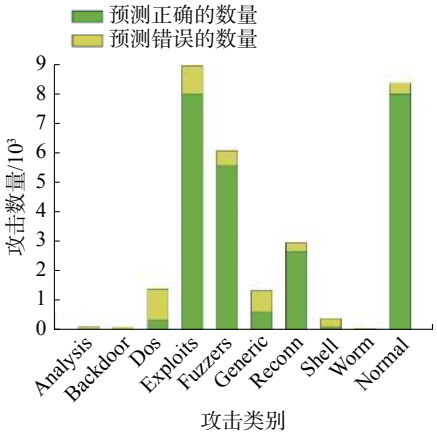

图 7　Netflow 数据源预测结果

Fig. 7　Results of Netflow data source

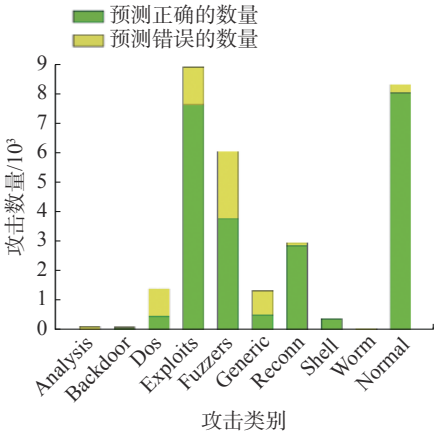

图 8　Snort 数据源预测结果

Fig. 8　Results of Snort data source

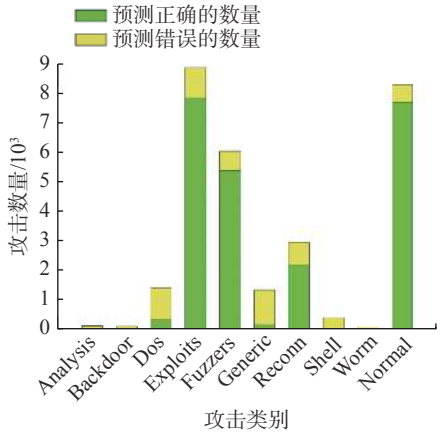

图 9　Suricata 数据源预测结果
Fig. 9　Results of Suricata data source

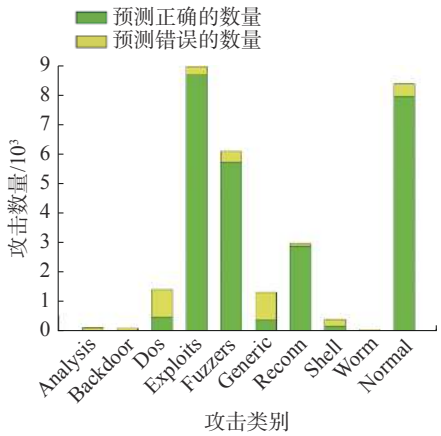

图 10　D-S 证据融合预测结果
Fig. 10　Attack prediction of D-S

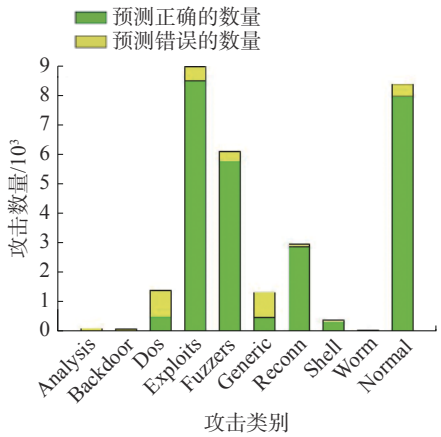

图 11　PSO-DS 证据融合预测结果
Fig. 11　Attack prediction of PSO-DS

如表 8 所示, 与其他结果相比, 本文选用的 PSO-DS 融合算法准确率高, 并且误警率降低, 充分证明了多点融合体系有效集成各单点检测的优势, 显著提高识别各攻击类型的能力; 针对误警率较高的问题, 本文查阅了使用此数据集的研究成果, 如表 9 所示, 均存在误警率较高问题。

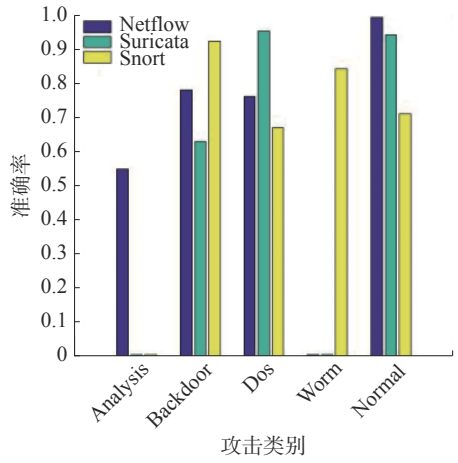

图 12　攻击预测
Fig. 12　Attack prediction

表 8　性能指标对比分析
Table 8　Performance index contrastive analysis　%

性能指标	Netflow	Snort	Surica-ta	D-S	PSO-DS
准确率	85.2	80.2	79.5	87.8	88.7
误警率	19.9	26.1	25.6	14.9	13.7

表 9　UNSW-NB15 已有成果实验结果[17]
Table 9　Other experimental results of UNSW-NB15[17]　%

性能指标	DT	LR	NB	ANN	EM clustering
准确率	85.6	83.2	82.1	81.3	78.5
误警率	15.8	18.5	18.6	21.1	23.8

2.2　网络安全态势评估

2.2.1　服务层态势

在网络中重放了 UNSW-NB15 中 2015-2-17 日的部分流量数据, 该段时间持续 33 000 s, 每 5 min 为一个时间窗口, 共 110 个时间窗口, 如横坐标 10 对应第 10 个时间窗口的态势值。使用本文提出的网络攻击威胁划分原则及权系数理论得到攻击因子, 按照评估体系进行评估。某一主机上 DNS、HTTP、SMTP 3 种服务的态势情况如图 13 所示。

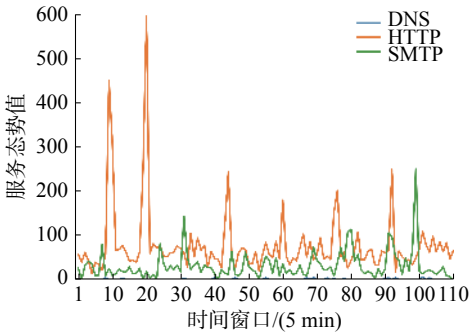

图 13　服务层安全态势
Fig. 13　Security situation of service

该日此主机上 HTTP 服务遭受到了 2 次强烈的攻击、4 次中度攻击、多次轻度攻击，管理员应该注意该主机 HTTP 服务的使用情况、该主机是否在访问非法网站、是否受到 web 攻击，并采取相应对策。

2.2.2　主机层态势

主机层态势与运行在主机上服务态势及各个服务重要度相关，根据服务的用户数量和使用频率确定服务的权值。由图 14 可知，主机 1 受到了 2 次强烈的攻击，并受到了多次小规模攻击；主机 2 受到了 2 次强烈的攻击；主机 3 受到了 5 次强烈的网络攻击；网络管理员应该特别注意此两台主机的运行状况。

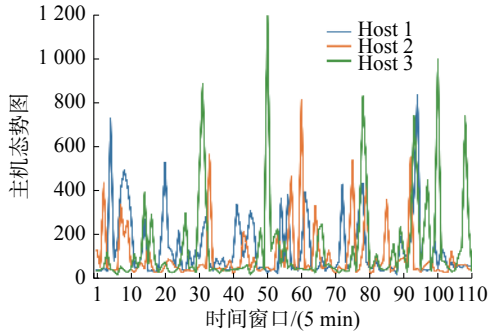

图 14　主机层安全态势

Fig. 14　Security situation of host

2.2.3　网络层态势

根据主机的用户数量与频率，确定主机的权值。根据主机的权值和主机的态势情况，计算得到整个网络的运行态势。由图 15 所示，该网络一天内持续受到攻击，产生 4 次较大波动，网络管理员应该查看这几个时间段内主机的运行情况，找到异常主机。此流量包一天内一直遭受到大量攻击，网络态势图符合流量包攻击情况，准确地展示出当天网络的运行状况。

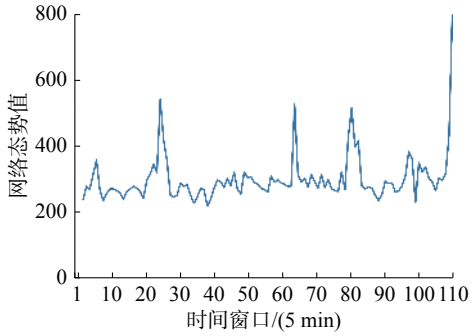

图 15　网络层安全态势

Fig. 15　Security situation of network

3　结束语

本文借鉴多源融合策略，充分利用层次化网络评估方法，提出了一个网络安全态势评估体系。1) 在网络中部署 Netflow、Snort、Suricata 探测器全面地获取网络数据；2) 基于恶意活动的特点，提炼有助于提高区别攻击类型的核心属性；3) 使用 BP 神经网络分别对数据进行决策；4) 利用 PSO-DS 方法对各 BP 神经网络分类结果进行有机融合；5) 使用层次化网络评估方法，直观展现网络态势。实验结果证明，不同探测器获得的网络信息对于识别不同攻击的优势不同，多点检测体系有机融合各单点检测的优势，显著提高识别各攻击类型的能力。

今后的研究将拓展到态势预测，通过态势曲线的变化趋势，对态势曲线的未来走向进行预测，以达到提前防护的目标。

参考文献：

[1] ENDSLEY MR. Toward a theory of situation awareness in dynamic system[J]. Human factors: the journal of the human factors and ergonomics society, 1995, 37(1): 32–6.

[2] BASS T. Intrusion detection systems and multisensor data fusion[J]. Communications of the ACM, 2000, 43(4): 99–105.

[3] 陈继军. 多传感器管理及信息融合 [D]. 西安: 西北工业大学, 2002, 49−55.
CHEN Jijun. Multi-Sensor administration and information fusion[D]. Xi'an: Northwestern Polytechnical University, 2002: 49–55.

[4] 诸葛建伟, 王大为, 陈昱, 等. 基于 D-S 证据理论的网络异常检测方法 [J]. 软件学报, 2006, 17(3): 463–471.
ZHUGE Jianwei, WANG Dawei, CHEN Yu, et al. A network anomaly detector based on the D-S evidence theory[J]. Journal of software, 2006, 17(3): 463–471.

[5] 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.
CHEN Xiuzhen, ZHENG Qinghua, GUAN Xiaohong, et al. A network anomaly detector based on the D-S evidence theory[J]. Journal of software, 2006, 17(4): 885–897.

[6] 马琳茹, 杨林, 王建新. 多源异构安全信息融合关联技术研究系 [J]. 系统仿真学报, 2008, 20(4): 981–989.
MA Linru, YANG Lin, WANG Jianxin. Research on security information fusion from multiple heterogeneous sensors[J]. Journal of system simulation, 2008, 20(4): 981–989.

[7] 韦勇, 连一峰, 冯登国. 基于信息融合的网络安全态势评估模型 [J]. 计算机研究与发展, 2009, 46(3): 353–362.
WEI Yong, LIAN Yifeng, FENG Dengguo. A network security situational awareness model based on information fusion[J]. Journal of computer research and development, 2009, 46(3): 353–362.

[8] 刘效武, 王慧强, 吕宏武, 等. 网络安全态势认知融合感控模型 [J]. 软件学报, 2016, 27(8): 2099–2114.

LIU Xiaowu, WANG Huiqiang, LU Hongwu, et al. Fusion-based cognitive awareness-control model for network security situation[J]. Journal of software, 2016, 27(8): 2099–2114.

[9] WANG Huan, CHEN Zhanfang, FENG Xin, et al. Research on network security situation assessment and quantification method based on analytic hierarchy process[J]. Wireless personal communications, 2018, 102(2): 1401–1420.

[10] 龚俭, 臧小冬, 苏琪, 等. 网络安全态势感知综述 [J]. 软件学报, 2017, 28(4): 1010–1026.
GONG Jian, ZANG Xiaodong, SU Qi, et al. Survey of network security situation awareness[J]. Journal of software, 2017, 28(4): 1010–1026.

[11] ZHAO Dongmei, LIU Jinxing. Study on network security situation awareness based on particle swarm optimization algorithm[J]. Computers and industrial engineering, 2018, 125: 764–775.

[12] 陈维鹏, 敖志刚, 郭杰, 等. 基于改进的 BP 神经网络的网络空间态势感知系统安全评估 [J]. 计算机科学, 2018, 45(11A): 345–347, 341.
CHEN Weipeng, AO Zhigang, GUO Jie, et al. Research on cyberspace situation awareness security assessment based on improved BP neural network[J]. Computer science, 2018, 45(11A): 345–347, 341.

[13] 贾焰, 韩伟红, 杨行. 网络安全态势感知研究现状与发展趋势 [J]. 广州大学学报(自然科学版), 2019, 18(3): 1–10.
JIA Yan, HAN Weihong, YANG Xing. Summary of network security situation assessment[J]. Journal of Guangzhou University (natural science edition), 2019, 18(3): 1–10.

[14] XI Rongrong, YUN Xiaochun, HAO Zhiyu. Framework for risk assessment in cyber situational awareness[J]. Iet information security, 2019, 13(2): 149–156.

[15] ZHENG Weifa. Research on situation awareness of network security assessment based on dempster-shafer[C]//2019 International Conference on Computer Science Communication and Network Security. France: Edition Diffusion Press Sciences, 2020: 131–136.

[16] MOUSTAFA N, SLAY J. UNSW-NB15: a comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set)[C]//Proceedings of 2015 Military Communications and Information Systems Conference. Canberra, Australia: IEEE, 2015: 1–6.

[17] MOUSTAFA N, SLAY J. The evaluation of network anomaly detection systems: statistical analysis of the UNSW-NB15 data set and the comparison with the KDD99 data set[J]. Information security journal: a global perspective, 2016, 25(1/2/3): 18–31.

[18] MOUSTAFA N, SLAY J, CREECH G. Novel geometric area analysis technique for anomaly detection using trapezoidal area estimation on large-scale networks[J]. IEEE transactions on big data, 2019, 5(4): 481–494.

[19] MOUSTAFA N, CREECH G, SLAY J. Big data analytics for intrusion detection system: statistical decision-making using finite dirichlet mixture models[M]. CARRASCOSA I P, KALUTARAGE H K, HUANG Yan. Data Analytics and Decision Support for Cybersecurity. Cham: Springer, 2017: 127-156.

[20] 甘文道, 周城, 宋波. 基于 RAN-RBF 神经网络的网络安全态势预测模型 [J]. 计算机科学, 2016,43 (11A): 388–392.
GAN Wendao, ZHOU Cheng, SONG Bo. Network security situation prediction model based on RAN-RBF neural network[J]. Computer science, 2016,43 (11A): 388–392.

[21] HECHT-NIELSEN R. Theory of the backpropagation neural network[C]//Proceedings of the International 1989 Joint Conference on Neural Networks. Washington, USA: IEEE, 1989: 593–605.

作者简介:

常利伟, 副教授, 中国计算机学会会员、中国密码学会会员、山西省区块链研究会理事, 主要研究方向为密码算法、网络安全态势感知、量子保密通信和区块链。参与国家级项目 4 项, 主持山西省科研及教研项目 3 项, 获山西省教学成果一等奖 1 项。发表学术论文近 20 篇。

田晓雄, 硕士研究生, 主要研究方向为网络安全和信息融合。

张宇青, 硕士研究生, 主要研究方向为网络安全与模式识别。