文章编号：2095-4980(2020)06-1093-06

WSN 中数据融合阶段的自适应入侵检测方案

李虹飞，申玉霞

(济源职业技术学院 教务处，河南 济源 459000)

摘　要：针对基于无线传感网络(WSN)的关键基础设施安全监测问题，提出一种基于数据融合阶段的自适应入侵检测算法。该算法以基于权重的簇化网络结构为基础，利用异常检测子系统和误用检测子系统分别检测已知攻击和未知攻击，然后通过跟踪 2 个子系统接收操作特征(ROC)和奖惩机制，自动调整转发至 2 个子系统的融合数据比例，即可实现在数据融合阶段对关键基础设施的自适应入侵检测。仿真分析表明：该算法的准确率和检测率高达 99.6%和 94.9%以上，与其他经典入侵检测系统相比，可分别至少提高 0.5%和 10.2%左右。

关键词：无线传感网络；数据融合；入侵检测；接收操作特征

中图分类号：TN911.7　　**文献标志码：**A　　**doi：**10.11805/TKYDA2020139

Adaptive intrusion detection scheme in data fusion stage of WSN

LI Hongfei, SHEN Yuxia

(Teaching Affairs Office, Jiyuan Vocational and Technical College, Jiyuan Henan 459000, China)

Abstract: In order to resolve the security monitoring problem of critical infrastructure based on Wireless Sensor Network(WSN), an adaptive intrusion detection scheme in data fusion stage is presented. The presented scheme takes the cluster network structure based on weight as the foundation, respectively using the anomaly detection subsystem and the misuse detection subsystem to detect the known attacks and the unknown attacks, and then adaptively adjusts the proportion of fusion data transmitted to the two sub-systems by tracking the Receiver Operating Characteristics(ROC) and the reward and punishment mechanism of the two sub-systems, so the intrusion detection of critical infrastructure in the data fusion stage can be realized. The simulation analysis indicates that the accuracy rate and detection rate of the presented algorithm is as high as 99.6% and 94.9% respectively, and can be improved at least 0.5% and 10.2% by comparing with other classic intrusion detection systems based on cluster structure.

Keywords: Wireless Sensor Networks; data fusion; intrusion detection; Receiver Operating Characteristic

随着无线传感网络(WSN)的广泛使用，其已成为监测如智能电网[1]、铁路/公路[2]、生产/制造[3]等关键基础设施的重要手段。而异常流量的检测活动，如入侵，对于关键基础设施的长期连续监测起着至关重要的作用。当前常用的入侵检测方案有数据挖掘、博弈理论、流量预测、计算智能等。文献[4]采用免疫理论中的阴性选择算法实现对 WSN 入侵活动的监测。文献[5]采用 K-means 聚类算法对 WSN 所采集的大量数据进行训练及分类来检测各类入侵攻击。文献[6]提出一种基于博弈论的入侵检测系统，用于有效检测 WSN 中的异常数据流。文献[7]提出一种基于簇结构的混合入侵检测系统(Clustered Hierarchical Hybrid-Intrusion Detection System，CHH-IDS)，该方案可以同时检测已知攻击和未知攻击，同时分析了数据流对入侵检测准确率的影响。文献[8]采用基于核自组织映射和粒子群优化(Kernel Self-Organizing Map and Particle Warm Optimization，KSOM-PSO)的神经网络对 WSN 进行入侵检测，可大幅提高入侵检测精确度。文献[9]提出一种 WSN 多阶段动态入侵检测博弈模型，主要利用贝叶斯规则修正下一阶段节点的后验概率，可明显提高簇形 WSN 的入侵检测准确率。文献[10]提出一种结合信任

收稿日期：2020-04-15；**修回日期：**2020-10-11
基金项目：国家自然科学基金资助项目(61371038)；近程高速目标探测技术国防重点学科实验室开放基金资助项目(30918014106)
作者简介：李虹飞(1975-)，女，副教授，主要研究方向为电子技术、自动控制。email:sheweic0703@163.com

机制和噪声检测技术的多协议层入侵检测(Trust-based Multi-Protocol layer Noise Intrusion Detection，T-MPNID)方法，用以解决跨层攻击检测中的高误报率和漏报率。目前多数入侵检测算法主要基于数据挖掘算法[11]，采用该类方案的 WSN 监测系统具有较好的入侵检测性能，但仍易遭受各类潜在的安全威胁和未知攻击，需要构建有效的入侵检测系统(Intrusion Detection System，IDS)，避免所采集的数据受已知和未知攻击，即通过检测异常活动提高 WSN 安全是建立 IDS 的根本目的[12-13]。识别 WSN 中异常活动的有效策略有机器学习、模糊逻辑、人工神经网络等。IDS 作为一种主动防御计算，主要通过二值分类区分正常行为和入侵行为，其中实施入侵检测的关键在于如何决定子系统的融合数据流。

因此，针对基于 WSN 的关键基础设施监测系统，通过分析 WSN 在感测数据融合阶段已知和未知的入侵行为，本文提出一种 WSN 数据融合阶段的自适应入侵检测方案(Adaptive Intrusion Detection scheme in Data Fusion Stage，AID-DFS)。方案首先利用两类机器学习子系统对数据进行处理：a) 异常检测子系统(Anomaly Detection Subsystem，ADS)，主要用于检测未知攻击，其基本原理是利用优化的 DBSCAN 分类器根据训练数据先获取非攻击模型，再比较感测数据来识别未知攻击；b) 误用检测子系统(Misuse Detection Subsystem，MDS)，主要用于检测已知攻击，其基本原理是利用随机森林分类器根据训练数据先获得攻击模型，再利用未来感测流量识别入侵行为。通过连续跟踪 2 个子系统的接收操作特征(ROC)以及结合 ROC 奖惩机制，从而实现自动调整转发至每个子系统的融合数据比例。与其他经典入侵检测方案相比，AID-DFS 方案的入侵检测准确率可高达 99.6%，入侵检测率可高达 94.9%。

1 网络模型

AID-DFS 方案采用簇化网络结构，见图 1。假设无线传感网络被划分成 n 个簇，每个簇包含 c 个节点，则一个簇的节点集可表示为：$\underset{(i\in1,2,\cdots,n)}{\text{Cluster}}\, i=\left\{CH_i\middle|S_j, j\in 0,1,\cdots,c-1\right\}$。各簇被选举的簇头(Cluster Head，CH)负责将簇内各成员节点转发的数据进行融合，然后再将融合后的数据转发至中心服务器(Center Server，CS)。

AID-DFS 方案的簇头选举方法采用基于权重的簇头产生机制[6]，即首先按照式(1)给出的权重计算方法，簇内每个节点都计算出其各自的簇头权重，然后再比较各节点的权重大小，最后选举权重最低的节点作为该簇的簇头。

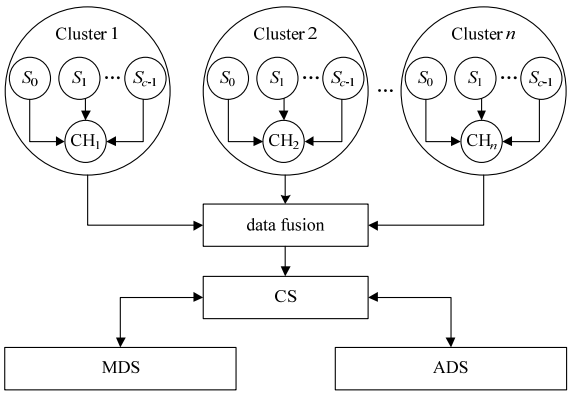

Fig.1 Network model of the AID-DFS scheme
图 1 AID-DFS 方案网络模型

$$W_j=\omega_1\Delta j+\frac{\omega_2}{\left|1/SRSS_j\right|}+\omega_3M_j+\omega_4\tau_j \tag{1}$$

式中：W_j 为节点 S_j 的权重；$\omega_1,\omega_2,\omega_3$ 与 ω_4 为式中各项的权重系数；$\Delta j=\left|d_j-m\right|$，$d_i$ 为 S_j 的度数(即 S_j 的一跳邻居节点数)，m 为单个簇头可处理的节点数；$SRSS_j$ 为 S_j 的接收信号强度，$\left|1/SRSS_j\right|$ 为 $SRSS_j$ 的归一化处理；M_j 为 S_j 的移动因子；τ_j 为 S_j 被选举为簇头的时间。

各簇选举出簇头后，簇头将簇内各成员节点的数据进行融合，并把融合后的数据传递至基站。AID-DFS 采用文献[7]中的融合算法计算融合节点的信任分，如式(2)所示：

$$T_{\text{agg}}=\frac{\left(\sum_{j=0}^{c-1}\left(T_j+1\right)\times T_{\text{agg}}^{j}\right)}{\sum_{j=0}^{c-1}\left(T_j+1\right)} \tag{2}$$

式中：T_j 为 S_j 的信任值；T_{agg}^{j} 为融合节点与 S_j 间的信任值。

2 AID-DFS 方案

AID-DFS 主要思路：首先，分别计算异常检测子系统和误用检测子系统在任意时刻的 ROC；然后，跟踪 2 个子系统的 ROC 变化，计算出决策函数；最后，根据决策函数实现动态调整向 2 个子系统转发融合数据的比例。

1) 分别确定 2 个子系统在 t_i 时刻的 ROC(即真假率)。其中，真假率的定义为真阳率(True Positive，TP)与假阳率(False Positive，FP)之比，如式(3)所示：

$$\begin{cases} M_{\mathrm{ADS}}(t_i)=\dfrac{P_{\mathrm{T,ADS}}(t_i)}{P_{\mathrm{F,ADS}}(t_i)} \\ M_{\mathrm{MDS}}(t_i)=\dfrac{P_{\mathrm{T,MDS}}(t_i)}{P_{\mathrm{F,MDS}}(t_i)} \end{cases} \tag{3}$$

式中：$M_{\mathrm{ADS}}(t_i)$ 为异常检测子系统在 t_i 时刻的接收操作特征；$M_{\mathrm{MDS}}(t_i)$ 为误用检测子系统在 t_i 时刻的接收操作特征。AID-DFS 为实施操作系统，可跟踪时间段 Δt 的变化值来确定 2 个子系统的操作特征值，如式(4)所示：

$$\begin{cases} M_{\mathrm{ADS}}(\Delta t)=\dfrac{P_{\mathrm{T,ADS}}(\Delta t)}{P_{\mathrm{F,ADS}}(\Delta t)} \\ M_{\mathrm{MDS}}(\Delta t)=\dfrac{P_{\mathrm{T,MDS}}(\Delta t)}{P_{\mathrm{F,MDS}}(\Delta t)} \end{cases} \tag{4}$$

式中 $\Delta t=t_{i+1}-t_i$。将 2 个子系统在 t_i 时刻和时间段 Δt 的接收操作特征进行加权融合整理后，可计算出 2 个子系统在 t_{i+1} 时刻的接收操作特征，如式(5)所示：

$$\begin{cases} M_{\mathrm{ADS}}(t_{i+1})=\alpha M_{\mathrm{ADS}}(t_i)+(1-\alpha)M_{\mathrm{ADS}}(\Delta t) \\ M_{\mathrm{MDS}}(t_{i+1})=\alpha M_{\mathrm{MDS}}(t_i)+(1-\alpha)M_{\mathrm{MDS}}(\Delta t) \end{cases} \tag{5}$$

式中 α 为权重参数。

2) AID-DFS 跟踪异常检测子系统和误用检测子系统中的接收操作特征，计算出任意时刻 t_i 的决策函数 $I(t_i)$，用于决策向 2 个子系统转发数据的比例，如式(6)所示：

$$I(t_i)=\frac{M_{\mathrm{ADS}}(t_i)}{M_{\mathrm{MDS}}(t_i)} \tag{6}$$

如果 $I(t_{i+1})>I(t_i)$，说明异常检测子系统优先于误用检测子系统，此时应增加向异常检测子系统转发融合数据的比例，即增加 $M_{\mathrm{ADS}}(t_i)$ 的感测数据比例，减少 $M_{\mathrm{MDS}}(t_i)$ 的感测数据比例：

$$\begin{cases} R_{\mathrm{a}}(t_{i+1})=R_{\mathrm{a}}(t_i)+\Delta R \\ R_{\mathrm{m}}(t_{i+1})=R_{\mathrm{m}}(t_i)-\Delta R \end{cases} \tag{7}$$

反之，则说明误用检测子系统优先于异常检测子系统，此时应增加向误用检测子系统转发融合数据的比例。即增加 $M_{\mathrm{MDS}}(t_i)$ 的融合数据比例，减少 $M_{\mathrm{ADS}}(t_i)$ 的融合数据比例：

$$\begin{cases} R_{\mathrm{a}}(t_{i+1})=R_{\mathrm{a}}(t_i)-\Delta R \\ R_{\mathrm{m}}(t_{i+1})=R_{\mathrm{m}}(t_i)+\Delta R \end{cases} \tag{8}$$

式中：$R_{\mathrm{a}}(t_i)$ 为 t_i 时刻向异常检测子系统转发融合数据的比例；$R_{\mathrm{m}}(t_i)$ 为 t_i 时刻向误用检测子系统转发融合数据的比例；ΔR 为转发融合数据的调整比例。AID-DFS 入侵检测流程如图 2 所示。

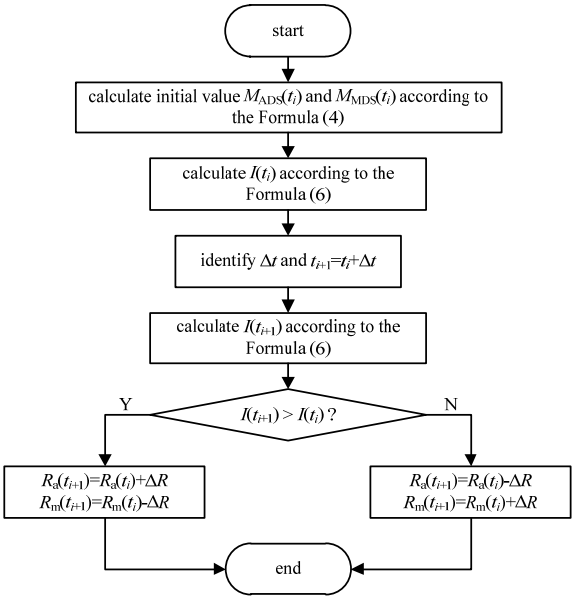

Fig.2 Intrusion detection flow of the AID-DFS

图 2 AID-DFS 方案入侵检测流程图

表 1 仿真参数

Table1 Simulation parameters

parameter	value
simulation time/s	660
communication radius/m	100
packet size/B	250
scope of trust score	[0,1]
weight parameter α	0.7
types of intrusion attack	DoS/Probing/U2R/R2L

3　性能分析

AID-DFS 采用 NS3 仿真器进行仿真验证。仿真无线传感器网络包含 25 个节点，随机部署在 100 m×100 m 的区域内，网络被划分为 4 个簇，节点间的通信采用层次—动态源路由协议。采用 KDD CUP 1999 数据库对 AID-DFS、CHH-IDS、KSOM-PSO 和 T-MPNID 进行性能分析比较。仿真无线传感器网络的其他参数如表 1 所示。

3.1 准确率

准确率(Accuracy Rate，AR)是指能够准确区分不同入侵攻击类型的概率：

$$R_{\mathrm{A}}=\frac{P_{\mathrm{T}}+N_{\mathrm{T}}}{P_{\mathrm{T}}+N_{\mathrm{T}}+P_{\mathrm{F}}+N_{\mathrm{F}}} \tag{9}$$

式中：N_{F}为假阴率；N_{T}为真阴率。

图 3 给出了随着入侵率的增加，AID-DFS,CHH-IDS,KSOM-PSO 与 T-MPNID 方案检测入侵类型的准确率变化情况，其中 ΔR=0.25。由图 3 可以看出，随着网络入侵率的增加，4 种方案检测入侵类型的准确率均呈现下降趋势，CHH-IDS 检测入侵类型的准确率下降幅度最快。当入侵率为 20%时，CHH-IDS 检测入侵类型的准确率为 99.7%，KSOM-PSO 检测入侵类型的准确率为 99.76%，T-MPNID 检测入侵类型的准确率为 99.83%，而 AID-DFS 方案检测入侵类型的准确率可高达 99.9%；当入侵率为 50%时，CHH-IDS 检测入侵类型的准确率已经降到 99.1%，KSOM-PSO 检测入侵类型的准确率为 99.23%，T-MPNID 方案检测入侵类型的准确率为 99.42%，而 AID-DFS 检测入侵类型的准确率仍可高达 99.6%。

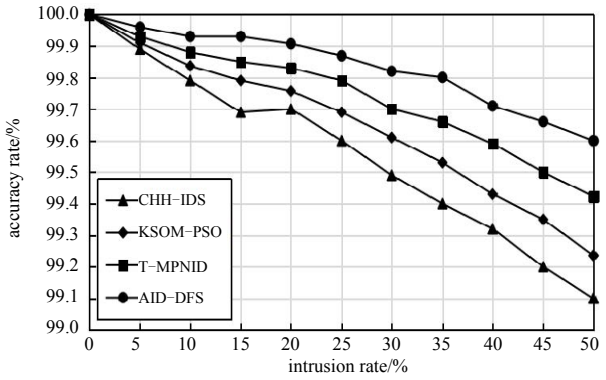

Fig.3 Comparison of intrusion detection accuracy among four schemes
图 3 4 种方案的入侵检测准确率比较

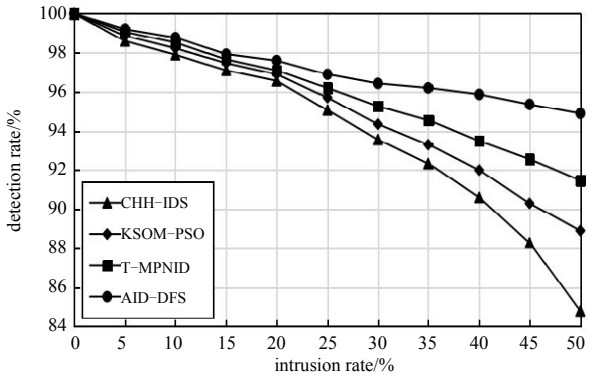

Fig.4 Comparison of intrusion detection rate among four schemes
图 4 4 种方案的入侵检测率比较

3.2 检测率

检测率(Detection Rate，DR)是指能够正确检测出入侵的概率：

$$R_{\mathrm{D}}=\frac{P_{\mathrm{T}}}{P_{\mathrm{T}}+P_{F}} \tag{10}$$

图 4 给出了随着入侵率的增加，AID-DFS,CHH-IDS,KSOM-PSO 与 T-MPNID 的检测率变化情况。由图 4 可以看出，随着入侵率的增加，4 种方案的检测率均呈现下降趋势，其中 CHH-IDS 检测率下降幅度更快。当入侵率为 30%时，CHH-IDS 的检测率为 93.6%，KSOM-PSO 的检测率为 94.4%，T-MPNID 的检测率为 95.3%，而 AID-DFS 方案的检测率可达到 96.5%；当入侵率为 50%时，CHH-IDS 的检测率为 84.7%，KSOM-PSO 的检测率为 88.9%，T-MPNID 的检测率为 91.5%，而 AID-DFS 方案的检测率仍可达 94.9%。

3.3 ROC 特性

图 5 给出了在不同转发调整比例 ΔR 下，AID-DFS 入侵检测方案的 TP 随 FP 的变化而变化的情况，该性能主要用于分析 AID-DFS 入侵检测方案的 ROC 特性。由图 5 可以看出，在不同的调整比例 ΔR 下，AID-DFS 入侵检测方案的 TP 随 FP 的增加均呈现上升趋势。当 ΔR=0 时，AID-DFS 入侵检测方案的 TP 最低；当 ΔR=0.25 时，AID-DFS 入侵检测方案的 TP 最高。由此可知，当 ΔR=0.25 时，AID-DFS 入侵检测方案可获得最优的 ROC 特性。

3.4 精确率曲线

精确率主要是指随查全率(Recall)变化，入侵检测方案的查准率(Precision)的变化情况。入侵检测方案的精确率越高，其性能越好，即当入侵检测方案的精确率趋于 1 时，表明该入侵方案的性能最好。其中，查全率的定义为：$P_{\mathrm{T}}/(P_{\mathrm{T}}+N_{\mathrm{F}})$，查准率的定义为：$P_{\mathrm{T}}/(P_{\mathrm{T}}+P_{\mathrm{F}})$。图 6 给出了不同转发调整比例 ΔR 下，AID-DFS 的精确率曲线变化情况。由图 6 可以看出，当 ΔR=0 时，AID-DFS 入侵检测方案的精确率最低；当 ΔR=0.25 时，AID-DFS

入侵检测方案的精确率最高。由此可知，$\Delta R=0.25$ 时，AID-DFS 入侵检测方案可以使系统性能达到最优，此时 AID-DFS 入侵检测方案的查全率为 99.8%，查准率可达 90.1%。

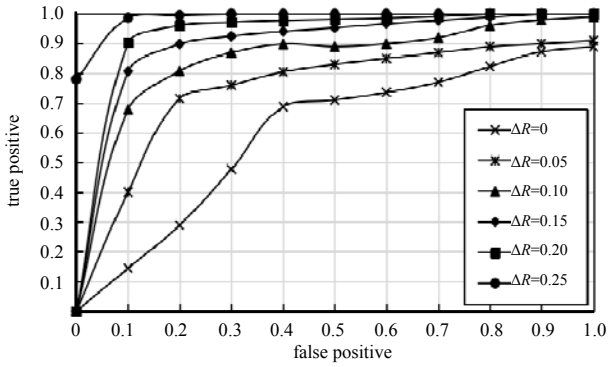

Fig.5 ROC characteristics under different ΔR
图 5 不同 ΔR 下的 ROC 特性

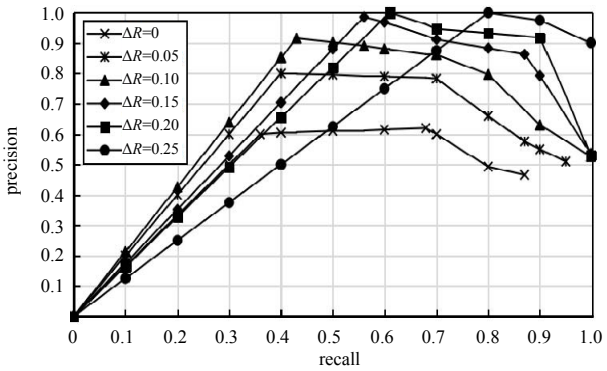

Fig.6 Accuracy curves under different ΔR
图 6 不同 ΔR 下的精确率曲线

4　结论

WSN 当前广泛用于监测关键基础设施。为有效提升 WSN 入侵检测的性能，给出了一种 WSN 数据融合阶段的自适应入侵检测方案。方案采用基于权重的簇化 WSN 结构，通过跟踪异常检测子系统和误用检测子系统的接收操作特征，再根据决策函数调整转发至 2 个子系统的数据比例，从而有效优化 WSN 的入侵检测性能。仿真数据表明，所给方案的准确率和检测率分别高达 99.6%和 94.9%以上，比其他经典入侵检测方案分别提高了约 0.5%和 10.2%，且实验确定当 $\Delta R=0.25$ 时，所给方案可获得最优接收操作特征性能，此时所给方案的查全率和查准率可分别达到 99.8%和 90.1%。

参考文献：

[1] OTOUM S,KANTARCI B,MOUFTAH H T. Hierarchical trust-based black-hole detection in WSN based smart grid monitoring[C]// 2017 IEEE International Conference on Communications(ICC). Paris,France:IEEE, 2017:1-6.

[2] 苏雪,张小云. 基于无线传感网络的高速铁路监测系统[J]. 太赫兹科学与电子信息学报, 2019,17(2):239-242. (SU Xue,ZHANG Xiaoyun. High-speed railway monitoring system based on wireless sensor network[J]. Journal of Terahertz Science and Electronic Information Technology, 2019,17(2):239-242.)

[3] 陈君,周达夫,王飞,等. 麦冬田间环境监测和自动灌溉系统设计与实现[J]. 太赫兹科学与电子信息学报, 2016,14(2):270-275. (CHEN Jun,ZHOU Dafu,WANG Fei,et al. Design and implementation of field monitoring and automatic irrigation system for radix ophiopogonis[J]. Journal of Terahertz Science and Electronic Information Technology, 2016,14(2):270-275.)

[4] RIZWAN R,KHAN F A,ABBAS H,et al. Anomaly detection in wireless sensor networks using immune-based bioinspired mechanism[J]. International Journal of Distributed Sensor Networks, 2015(6):1120-1127.

[5] WAZID M,DAS A K. An efficient hybrid anomaly detection scheme using K-means clustering for wireless sensor networks[J]. Wireless Personal Communications, 2016,90(4):1971-2000.

[6] SEDJELMACI H,SENOUCI S M,TALEB T. An accurate security game for low-resource IoT devices[J]. IEEE Transactions on Vehicular Technology, 2017,66(10):9381-9393.

[7] OTOUM S,KANTARCI B,MOUFTAH H T. Detection of known and unknown intrusive sensor behavior in critical applications[J]. IEEE Sensors Letters, 2017,1(5):1-4.

[8] 刘双,石飞,汪烈军,等. 基于 KSOM-PSO 算法的无线传感器网络入侵检测研究[J]. 中国科技论文, 2017,12(2):148-153. (LIU Shuang,SHI Fei,WANG Liejun,et al. Research on intrusion detection in wireless sensor networks based on KSOM-PSO algorithm[J]. China Science Paper, 2017,12(2):148-153.)

[9] 周伟伟,郁滨. WSNs 多阶段入侵检测博弈最优策略研究[J]. 电子与信息学报, 2018,40(1):63-71. (ZHOU Weiwei,YU Bin. Optimal defense strategy in WSNs based on the game of multi-stage intrusion detection[J]. Journal of Electronics & Information Technology, 2018,40(1):63-71.)

(下转第 1127 页)