**引文格式**：苏明. 无线传感网络中的自适应入侵检测算法[J]. 导航定位学报, 2020, 8(4): 106-110.（SU Ming.Adaptive intrusion detection in wireless sensor networks[J]. Journal of Navigation and Positioning, 2020, 8(4): 106-110.）DOI:10.16547/j.cnki.10-1096.20200418.

# 无线传感网络中的自适应入侵检测算法

苏　明

（北京开放大学，北京　100081）

**摘要**：针对无线传感网络的入侵检测算法因大多基于数据挖掘算法，所建立的监测系统易遭受未知攻击的问题，提出 1 种自适应的入侵检测算法：跟踪每个子系统的接收操作特征（ROC），再依据 ROC 行为特征，自适应地调整转发至每个子系统的融合数据的比例，达到在数据融合阶段运行入侵检测系统，实现有效监测的目的。仿真结果表明，基于该算法的监测系统具有 99 %的准确率。

**关键词**：无线传感网络；入侵检测；机器学习；簇；接收操作特征

**中图分类号**：TPT393　　**文献标志码**：A　　**文章编号**：2095-4999(2020)05-0106-05

## Adaptive intrusion detection in wireless sensor networks

*SU Ming*

（Beijing Open University, Beijing 100018, China）

**Abstract**：Aiming at the problem that it is liable to unknown attack for the monitoring and detection system established by the intrusion detection algorithm of wireless sensor networks mostly based on data-mining method, the paper proposed an adaptive intrusion detection algorithm: the receiver operating characteristics (ROC) of each subsystem were tracked, and based on the improvement/degradation of the ROC behavior, the proportion of aggregated data of each subsystem forwarded was adaptivley adjusted, which leads to the efficient monitoring and detection by implementing the intrusion detection in the data fusion phase. Simulational result showed that the sytem with the proposed algorithm could provide up to an accuracy rate of 99 %.

**Keywords**：wireless sensor networks; intrusion detection; machine learning; clustering; receiver operating characteristics

## 0　引言

无线传感网络（wireless sensor networks, WSNs）已在多个领域内广泛使用[1-2]，如智能家居、智能电网等。WSNs 中的传感节点，能够感知应用环境中的异常事情，因此可以利用节点的感知能力，来检测异常事情，例如在智能家居的安防中，可以用红外传感节点感知异常物体入侵。

现存的多数入侵检测算法都依赖数据挖掘算法[3-5]。尽管它们在入侵检测方面有较好的性能，但是基于传感网络的监测系统，仍容易遭受网络攻击。因此，有效的入侵检测系统（intrusion detection system, IDS）非常关键，通过 IDS 可以避免数据受已知和未知攻击。

通过检测异常活动，提高网络安全[6]是部署 IDS 的根本目的。计算智能，包括机器学习、模糊逻辑、人工神经网络等均是识别网络流量中异常活动的有效策略。IDS 就是通过二值分类，区分正常行为和入侵行为。

文献[7]通过实时自适应模型产生（adaptive model generation, AMG）结构，实施基于数据挖掘的 IDS 系统。文献[8]对基于异常的 IDS 进行分析，提出基于博弈理论的入侵检测系统。

为此，本文针对基于 WSNs 应用的环境监测

**收稿日期**：2019-11-18

**基金项目**：北京市教委-自然科学基金重点项目（KZ201951160050）。

**作者简介**：苏明（1975—），女，辽宁沈阳人，硕士，副教授，研究方向为人工智能、计算机网络。

系统，分析了在感测数据融合阶段的已知和未知的入侵行为，然后提出自适应的入侵检测（adaptive intrusion detection, AID）系统。

在 AID 中，利用 2 类机器学习子系统对数据进行处理：①误用检测子系统（misuse detection Subsystem, MDS）；②异常检测子系统（anomaly detection subsystem, ADS）。MDS 能够有效地检测已知攻击，ADS 能够检测未知攻击。所谓已知攻击是系统已掌握了攻击特点的攻击；未知攻击是指系统对攻击特点并不了解的攻击。

MDS 通过随机森林分类器检测已知攻击。它先通过训练数据，获取攻击模型，然后利用未来感测流量，识别入侵行为。而 ADS 是通过优化的 DBSCAN 分类器，来检测未知攻击：先依据训练数据获取正常模型（非攻击模型），再利用模型识别未知攻击。

实施入侵检测的关键在于，如何决定子系统的融合数据流，例如文献[9]提出簇结构的混合入侵检测系统（clustered hierarchical hybrid-intrusion detection system, CHH-IDS），CHH-IDS 就是通过分析数据流对入侵检测准确率的影响，达到提高准确地检测入侵。本文提出的 AID 系统，可以连续地跟踪每个子系统的接收操作特征（receiver operating characteristics, ROC），再结合 ROC 的奖惩机制，自动调整给每个子系统转发的融合数据比例来完成入侵检测。

## 1　预备知识

### 1.1　基于权重函数的簇结构

引用簇化的网络结构。假定网络有 $N$ 个簇，每个簇由 $c$ 个传感节点。在每个簇内，簇头（cluster head, CH）负责融合簇内传感节点所转发的数据。一旦融合完毕，CH 就将数据转发至中心服务器，如图 1 所示。

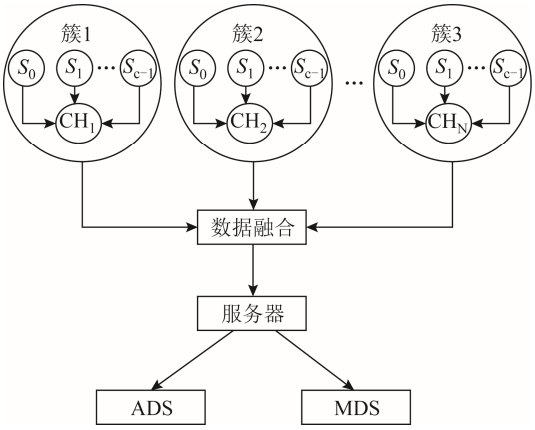


**图 1　系统模型**

AID 系统引用基于权重的簇头产生机制[10]，给每个节点定义 1 个簇头权重，具有最低权重的节点成为簇头。具体而言，令 $W_i$ 表示传感节点 $s_i$ 的权重，其定义为

$$W_i=\omega_1\Delta i+\frac{\omega_2}{\left|1/\mathrm{SRSS}_i\right|}+\omega_3 M_i+\omega_4\tau_i \tag{1}$$

式中：$\Delta i=\left|d_i-n\right|$，$d_i$ 为节点 $s_i$ 度数，即节点 $s_i$ 的一跳邻居节点数，$n$ 为 1 个簇头能够处理的节点数；$\mathrm{SRSS}_i$ 表示节点 $s_i$ 的接收信号强度值，$1/\mathrm{SRSS}_i$ 是对 $\mathrm{SRSS}_i$ 的归一化处理；$M_i$ 为传感节点的移动因子；$\tau_i$ 为成为簇头的时间；$\omega_1$、$\omega_2$、$\omega_3$ 和 $\omega_4$ 为各项变量的权重系数。每个节点估计自己的权重，并向它的邻居节点广播自己权重，最终将具有最低权重的节点作为簇头[10]。

### 1.2　数据融合

每个簇头融合其簇内传感节点的数据，然后将融合的数据传输至信宿。AID 系统引用文献[11]的数据融合算法。通过计算融合节点的信任值，其定义为

$$T_{\mathrm{agg}}=\frac{\left(\sum_{i=0}^{N-1}\left(T_i+1\right)\cdot T_{\mathrm{agg}}^{i}\right)}{\sum_{i=0}^{N-1}\left(T_i+1\right)} \tag{2}$$

式中：$T_{\mathrm{agg}}$ 为融合节点的信任值；$T_i$ 为 $s_i$ 的信任值；$T_{\mathrm{agg}}^{i}$ 为融合节点与节点 $s_i$ 间的信任值；$N$ 为簇内的传感节点数。

## 2　AID 系统

AID 系统旨在跟踪 MDS 和 ADS 子系统中 ROC 的变化，并调整向它们转发感测数据的比例[12]。引用真假率评估 ADS 和 MDS 检测入侵的性能，即：

$$M_1\left(t_i\right)=\frac{\mathrm{TP}_1\left(t_i\right)}{\mathrm{FP}_1\left(t_i\right)} \tag{3}$$

$$M_2\left(t_i\right)=\frac{\mathrm{TP}_2\left(t_i\right)}{\mathrm{FP}_2\left(t_i\right)} \tag{4}$$

式中：$M_1\left(t_i\right)$ 为在时刻 $t_i$，ADS 系统的真假率；$\mathrm{TP}_1\left(t_i\right)$ 为在时刻 $t_i$，ADS 系统将非入侵事件正确判断为非入侵事件的个数，即真阳（true positive, TP）次数；$\mathrm{FP}_1\left(t_i\right)$ 为在时刻 $t_i$，ADS 系统将非入侵事件正确判断为入侵事件的个数，即假阳（false positive, FP）次数；$M_2\left(t_i\right)$ 为在时刻 $t_i$，MDS 系统的真假

率；$\mathrm{TP}_2(t_i)$为在时刻$t_i$，MDS 系统将非入侵事件正确判断为非入侵事件的个数；$\mathrm{FP}_2(t_i)$为在时刻$t_i$，MDS 系统将非入侵事件正确判断为入侵事件的个数。

单一时刻点的真假率并不能准确地反映 AID 和 MDS 系统的检测入侵性能。为此，利用一段时间$\Delta t$观察真假率。AID 和 MDS 系统在$\Delta t$时间内的的真假率的计算方法为：

$$M_1(\Delta t)=\frac{\mathrm{TP}_1(\Delta t)}{\mathrm{FP}_1(\Delta t)} \tag{5}$$

$$M_2(\Delta t)=\frac{\mathrm{TP}_2(\Delta t)}{\mathrm{FP}_2(\Delta t)} \tag{6}$$

式中：$\Delta t=t_{i+1}-t_i$。

用真假率表述系统的 ROC。为了能准确地跟踪真假率的变化情况，下 1 个时刻的真假率包含当前时刻的真假率和时间段$\Delta t$内的真假率。AID 系统和 MDS 系统在时刻$t_{i+1}$的真假率的计算方法为：

$$M_1(t_{i+1})=\alpha M_1(t_i)+(1-\alpha)M_1(\Delta t) \tag{7}$$

$$M_2(t_{i+1})=\alpha M_2(t_i)+(1-\alpha)M_2(\Delta t) \tag{8}$$

式中：$t_{i+1}=t_i+\Delta t$；$\alpha$为权重参数。

除了每个子系统的 ROC 行为，AID 系统跟踪 2 个子系统（ADS 和 MDS）在任意时刻$t_i$的平均 ROC，即

$$I(t_i)=\frac{M_1(t_i)}{M_2(t_i)} \tag{9}$$

对于任意时刻$t_i$，如果$I(t_i)>I(t_{i-1})$，则表明 ADS 子系统优先 MDS 子系统，就增加向 ADS 转发感测数据的比例。相反，如果$I(t_i)<I(t_{i-1})$，MDS 子系统优先 ADS 子系统，就增加向 MDS 转发感测数据的比例。

具体而言，若如果$I(t_{i+1})>I(t_i)$，就增加 ADS 系统的感测数据比例、减少 MDS 系统的比例，即

$$\left.\begin{aligned}R_a(t_{i+1})&=R_a(t_i)+\Delta R\\R_m(t_{i+1})&=R_m(t_i)-\Delta R\end{aligned}\right\} \tag{10}$$

式中：$R_a(t_i)$为在时刻$t_i$向 ADSs 转发数据的比例；$R_m(t_i)$为在时刻$t_i$向 MDSs 转发的数据比例[13]；$\Delta R$为调整比例。整个过程如图 2 所示。

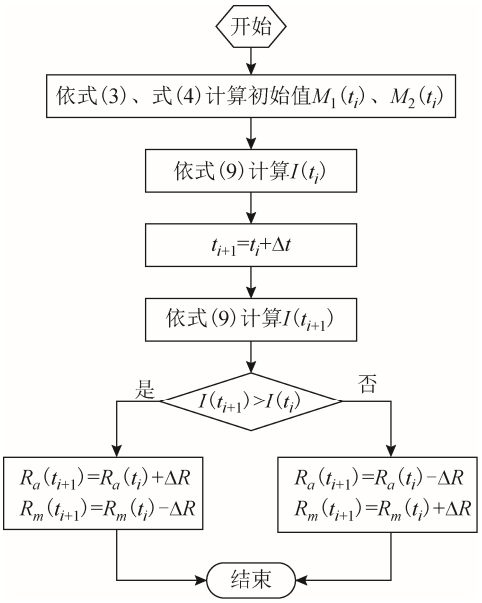


**图 2　AID 系统流程**

## 3　实验与结果分析

为了更好地分析 AID 系统，引用 NS3 仿真器建立仿真平台。在 100 m×100 m 区域部署 20 个传感节点，将这些传感节点分成 4 个簇。引用层次-动态源路由（hierarchical-dynamic source routing, HDSR）协议完成节点间通信。

引用数据挖掘的知识发现（knowledge discovery in data mining, KDD）CUP 1999 数据库，通过 KDD CUP 1999 数据库评估 AID 系统检测性能。并考虑 4 类攻击：否认服务（denial of service, DoS）、端口（probe）攻击、远程用户攻击（remote-to-login, R2L）、提权（user-to-root, U2R）攻击。具体的仿真参数如表 1 所示。

**表 1　仿真参数**

| 仿真参数 | 值 |
|---|---|
| 节点数 | 20 |
| 路由协议 | H-DSR |
| 簇数 | 4 |
| 数据包尺寸/bit | 250 |
| 信任分范围 | [0, 1] |
| 通信半径/m | 100 |
| 仿真区域 | 100 m×100 m |
| 攻击类型 | Dos,Probing, U2R, R2L |
| $\alpha$ | 0.7 |
| 仿真时间/s | 600 |

### 3.1　准确率

准确率 AR 表示分类的准确性，其定义为

$$\mathrm{AR}=\frac{N_{\mathrm{TP}}+N_{\mathrm{TN}}}{N_{\mathrm{TP}}+N_{\mathrm{TN}}+N_{\mathrm{FP}}+N_{\mathrm{FN}}} \tag{11}$$

式中：$N_{\mathrm{TP}}$ 表示将非入侵事件正确判断为非入侵事件的个数；$N_{\mathrm{TN}}$ 表示入侵事件正确判断为入侵事件的个数；$N_{\mathrm{FP}}$ 表示将非入侵事件错误判断为入侵事件的个数；$N_{\mathrm{FN}}$ 表示将入侵事件错误判断为非入侵事件的个数。

图 3 显示了 AR 随入侵率的变化情况，且 $\Delta R$=0.25。

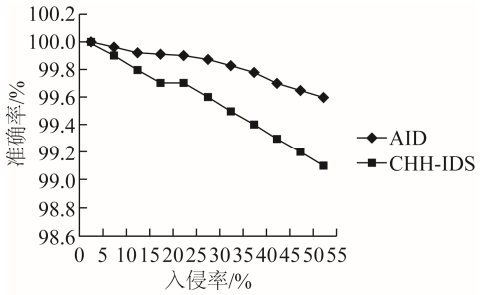


**图 3　准确率随入侵率的变化情况**

从图 3 可知，AR 随入侵率的增加而下降。原因在于，入侵率越高，入侵者越多，检测难度越高，降低了准确率。相比于 CHH-IDS，AID 系统提高了准确率，AID 系统的准确率达到 99 %以上。

### 3.2　检测率

AID 系统的检测率（detection rate, DR），反映了正确地检测入侵事件的概率，其定义为

$$\mathrm{DR}=\frac{N_{\mathrm{TP}}}{N_{\mathrm{TP}}+N_{\mathrm{FP}}} \tag{12}$$

图 4 显示了 DR 随入侵率的变化情况。

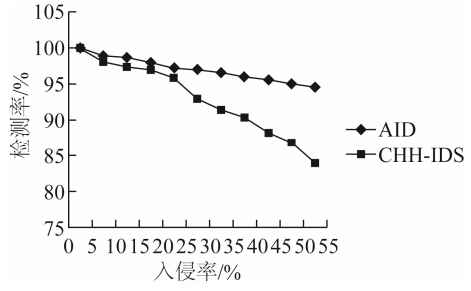


**图 4　检测率随入侵率的变化情况**

从图 4 可知，入侵率的增加，降低了 DR。但当入侵率为 50 %，AID 系统的 DR 仍达到 95 %，远高于 CHH-IDS。例如，当入侵率为 40 %，AID 系统的 DR 为 96 %，而 CHH-IDS 系统的 DR 只达到 88%。

### 3.3　TP 性能

图 5 显示了 TP 随 FP 的变化曲线，其反映了 ROC 特性。从图 5 可知：当 $\Delta R=0$ 时，TP 最低；而当 $\Delta R=0.25$ 时，TP 最高，即当 $\Delta R=0.25$ 时，能够获取最优的 ROC 特性。

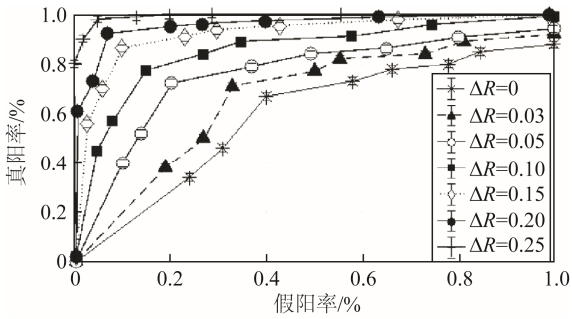


**图 5　TP 随 FP 的变化情况**

### 3.4　精确率曲线

精确率曲线反映了查准率（precision）随查全率（recall）的变化过程。其中查全率等于 $N_{\mathrm{TP}}/\left(N_{\mathrm{TP}}+N_{\mathrm{FN}}\right)$，而查准率等于 $N_{\mathrm{TP}}/\left(N_{\mathrm{TP}}+N_{\mathrm{FP}}\right)$。精确率越高（趋于 1），性能越好。

图 6 显示了 $\Delta R$ 值变化时的精确率曲线。

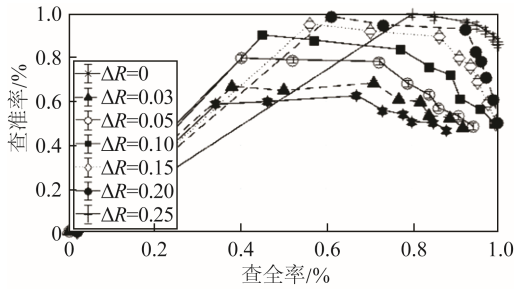


**图 6　精确率曲线**

从图 6 可知，当 $\Delta R=0.25$ 时，能够获取最高的精确率。具体而言，将 $\Delta R$ 设置为 0.25，查全率为 99.8 %、查准率为 90.1 %时，系统性能达到最优。

## 4　结束语

针对重要网络基础设施的监测问题，提出自适应检测 AID 系统。AID 系统以簇化的 WSNs 结构为基础，通过跟踪 ROC 特征，调整转发至 MDS 和 ADS 2 个子系统的数据的比例，进而优化系统。仿真数据表明，提出的 AID 系统能有效地提高了入侵检测率，其准确率可达到 99 %。

## 参考文献

[1]　ALCARAZ C, ZEADALLY S. Critical infrastructure protection: requirements and challenges for the 21st century[J]. International journal of critical infrastructure protection, 2015(8): 53-66.

[2] 刘秀平. 浅析当前网络入侵检测系统的方案研究[J]. 数码世界, 2016(4): 55-56.
[3] 刘强, 蔡志平, 殷建平,等. 网络安全检测框架与方法研究[J]. 计算机工程与科学, 2017, 39(12): 2224-2229.
[4] OTOUM S, KANTARCI B, MOUFTAH H T. Hierarchical trust-based black-hole detection in WSN-based smart grid monitoring[C]//The Institute of Electrical and Electronic Engineers(IEEE). Proceeding of 2017 IEEE International Conference on Communications (ICC). Paris: IEEE, 2017: 1-6.
[5] GRIMALDI S, GIDLUND M, LENNVALL T, et al. Detecting communication blackout in industrial wireless sensor networks[C]//The Institute of Electrical and Electronic Engineers(IEEE). Proceeding of 2016 IEEE World Conference on Factory Communication Systems (WFCS). Aveiro: IEEE, 2016: 1-8.
[6] JAYSHREE K. Intrusion detection using data mining approach[J]. International Journal of Science and Research (IJSR), 2014, 3(1): 1142-1145, 2014.
[7] HONIG A, HOWARD A, ESKIN E, et al. Adaptive model generation: an architecture for deployment of data mining-based intrusion detection systems[EB/OL]. [2019-08-28]. https: //academiccommons. columbia. edu/doi/10. 7916/D88D031B.
[8] SEDJELMACI H, SEDJELMACI S M, TALEB T. An accurate security game for low-resource iot devices[J]. IEEE Transactions on Vehicular Technology, 2017, 66(10): 9381-9393.
[9] 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.
[10] BELABED F, BOUALLEGUE R. An optimized weight-based clustering algorithm in wireless sensor networks[C]//The Institute of Electrical and Electronic Engineers(IEEE). Proceeding of 2016 International Wireless Communications and Mobile Computing Conference (IWCMC). Byblos, Lebanon: IEEE, 2016: 45-52.
[11] ZHANG W, DAS S K, LIU Y. A trust based framework for secure data aggregation in wireless sensor networks[C]//The Institute of Electrical and Electronic Engineers(IEEE). Proceeding of IEEE Communications Society on Sensor and Ad Hoc Communications and Networks. Reston, VA: IEEE, 2016: 34-42. DOI: 10. 1109/SAHCN. 2006. 288409.
[12] 钱亚冠, 卢红波, 纪守领. 一种针对基于 SVM 入侵检测系统的毒性攻击方法[J]. 电子学报, 2019, 47(1): 59-65.
[13] 王萌, 王亚刚, 韩俊刚. 基于 NDNN 的入侵检测系统[J]. 微电子学与计算机, 2018, 35(7): 89-92.

（上接第 67 页）

[11] CHEN Y, YI Z C. The BP artificial neural network model on expressway construction phase risk[J]. Systems Engineering Procedia, 2012, 4: 409-415.
[12] 陈阳, 胡伍生, 严宇翔, 等. 基于神经网络模型误差补偿技术的对流层延迟模型研究[J]. 大地测量与地球动力学, 2018, 38(6): 577-580, 586.
[13] 王德明, 王莉, 张广明. 基于遗传 BP 神经网络的短期风速预测模型[J]. 浙江大学学报(工学版), 2012, 46(5): 837-841, 904.
[14] 尹光志, 李铭辉, 李文璞, 等. 基于改进 BP 神经网络的煤体瓦斯渗透率预测模型[J]. 煤炭学报, 2013, 38(7): 1179-1184.
[15] BEVIS M, BUSINGER S, HERRING T, et al. GNSS meteorology-remote sensing of atmospheric water vapor using the global positioning system[J]. Journal of Geophysical Research, 1992, 97( D14) : 15787-15801.
[16] 于胜杰, 柳林涛. 水汽加权平均温度回归公式的验证与分析[J]. 武汉大学学报(信息科学版), 2009, 34(6): 741-744.
[17] 江婷, 李黎, 田莹, 等. 基于湖南 CORS 网的 PWV 时空变化分析及其在暴雨落区预报中的应用[J]. 大地测量与地球动力学, 2018, 38(7): 707-713.
[18] 姚宜斌, 郭健健, 张豹, 等. 湿延迟与可降水量转换系数的全球经验模型[J]. 武汉大学学报(信息科学版), 2016, 41(1): 45-51.
[19] SAASTAMOINEN J. Atmospheric correction for the troposphere and stratosphere in radio ranging satellites[J]. The Use of Artificial Satellites for Geodesy, 1972, 3(15): 247-251.