Journal of System Simulation

Volume 28 | Issue 6 Article 12

6-8-2020

Research of Intrusion Alert Aggregation Based on Spatial and Temporal Density

Zhang Jing
PLA Information Engineering University, Zhengzhou 450004, China;

Hengjun Wang
PLA Information Engineering University, Zhengzhou 450004, China;

Junquan Li
PLA Information Engineering University, Zhengzhou 450004, China;

Bin Yu
PLA Information Engineering University, Zhengzhou 450004, China;

Research of Intrusion Alert Aggregation Based on Spatial and Temporal Density

Abstract

Abstract: Distributed Intrusion Detection System has created the problem to investigate a mass of duplicate alerts and high false positive rate in practical applications. Based on DBSCAN, *density based spatial and temporal clustering of applications with noise (DBS&TCAN) algorithm was proposed by introducing temporal density. The approach aggregated partial alerts based on spatial density, and merges partial aggregation on the basis of temporal density.* The effectiveness of the algorithm was demonstrated by the intrusion detection evaluation dataset. The comparative experiments and analysis show that *the algorithm is effective in alert aggregation and gives better results in terms of real time*.

Keywords

Intrusion detection system, alert aggregation, temporal density, DBSCAN, DBS&TCAN;, real time

Recommended Citation

Zhang Jing, Wang Hengjun, Li Junquan, Yu Bin. Research of Intrusion Alert Aggregation Based on Spatial and Temporal Density[J]. Journal of System Simulation, 2016, 28(6): 1336-1343.

第 28 卷第 6 期 系统仿真学报© Vol. 28 No. 6
2016 年 6 月 Journal of System Simulation Jun., 2016

基于空间和时间密度的入侵报警聚合研究

张靖，王衡军，李俊全，郁滨
（解放军信息工程大学，河南 郑州 450004）

摘要：针对分布式入侵检测系统在实际应用中存在大量重复报警和高误报率的问题，在研究 DBSCAN 算法的基础上，引入时间密度，提出一种基于空间和时间密度的抗噪声聚合算法(DBS&TCAN)。基于空间密度聚合局部报警信息和时间密度对局部聚合结果进行合并，可以有效减少重复报警并降低误报率。实验采用数据集测试的方法对算法进行了测试，并与相关研究工作进行比较和分析。结果表明，该算法具有较好的聚合效果，并在实时性方面体现出优势。

关键词：入侵检测系统；报警聚合；时间密度；DBSCAN；DBS&TCAN；实时性
中图分类号：TP393.08 文献标识码：A 文章编号：1004-731X (2016) 06-1336-08

Research of Intrusion Alert Aggregation Based on Spatial and Temporal Density

Zhang Jing, Wang Hengjun, Li Junquan, Yu Bin
(PLA Information Engineering University, Zhengzhou 450004, China)

Abstract: Distributed Intrusion Detection System has created the problem to investigate a mass of duplicate alerts and high false positive rate in practical applications. Based on DBSCAN, *density based spatial and temporal clustering of applications with noise (DBS&TCAN) algorithm was proposed by introducing temporal density. The approach aggregated partial alerts based on spatial density, and merges partial aggregation on the basis of temporal density.* The effectiveness of the algorithm was demonstrated by the intrusion detection evaluation dataset. The comparative experiments and analysis show that *the algorithm is effective in alert aggregation and gives better results in terms of real time*.

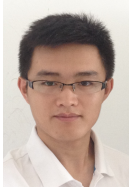
Keywords: Intrusion detection system; alert aggregation; temporal density; DBSCAN; DBS&TCAN; real time

引言

分布式入侵检测[1]可分为多点数据采集和多引擎协作入侵检测，前者基于相同的检测引擎，在多点进行数据采集，丰富了检测数据的来源；后者采用异构的检测引擎，可以弥补单一引擎的缺陷。以上两种分布式入侵检测技术，带来了相同的问题，即警报数量急剧增加，其中存在大量的重复报警和误报警，对管理员分析和管理这些报警信息造成了巨大的困难。

对报警进行聚合和关联是解决这一困难的有效办法[2]，通过报警聚合和关联，可以在保证较高检测率的前提下，得到尽量多与攻击相关的报警，有效减少重复报警并降低误报率。其中，报警聚合是将由同一安全事件诱发的大量性质相同或相近的报警合并成一个报警，也称报警聚类。

文献[3-4]将报警集合划分为 k 个子集，每个子集代表一个类别，当报警信息密集分布且类之间距离较远时聚合效果较好，但事先给定的聚类数对聚



收稿日期：2015-04-29 修回日期：2015-07-24;
作者简介：张靖(1991-)，男，安徽滁州，硕士生，研究方向为仿真、信息安全技术；王衡军（1973-），男，湖南衡阳，博士，副教授，研究方向为人工智能、信息安全；李俊全(1965-)，男，河北涿州，博士，研究员，博导，研究方向为密码学、信息安全。

合结果影响较大，且对噪声敏感。文献[5-6]采用统计方法[5]和神经网络[6]获取报警的聚合模型，充分考虑噪声数据，但该方法假设报警属性概率分布相互独立，且概率分布的更新和存储开销较大，不适用于大数据量的报警集合。而文献[7]对报警集合进行层次式的分解，能较好的应用于大数据量的数据集，但存在无法回溯，且合并点和分裂点选择困难的不足。文献[8]采用 CLIQUE 算法将报警集合划分为有限个网格单元，在各个网格中单独进行聚合操作，处理速度快，但聚类精度不高，影响报警的进一步处理。文献[9-10]基于密度的方法，将报警集合的高密度区域划分为一个类，能够发现任意形状的类，且聚类精度较高，但对输入参数敏感。

DBSCAN 算法[9]是基于空间密度的抗噪声聚类算法，该算法利用基于密度的聚合方法，将具有足够高密度的区域划分为簇。其显著优点是聚类速度快且能够在带有噪声的空间发现任意形状的聚类，可以很好的应用于报警聚合过程。然而，在实际应用中，DBSCAN 算法处理海量报警时，需要较大的 I/O 和内存开销，实时性差；同时，对输入参数敏感，当报警信息密度不均匀时，聚合效果较差，误报率高。

针对 DBSCAN 算法存在的不足，本文引入时间密度，提出一种基于空间和时间密度的抗噪声聚合算法(Density-Based Spatial and Time Clustering of Applications with Noise，DBS&TCAN)用于多引擎协作检测场合，能够有效减少重复报警并降低误报率。实验采用数据集测试的方法对算法进行了测试，并与相关研究工作进行比较。实验结果表明，该算法具有较好的聚合效果，并在实时性方面体现出优势。

1 报警聚合模型

报警聚合模型如图 1 所示，IDS 端，首先对异构 IDS 产生的原始报警信息格式进行标准化，再预处理后生成中间报警；而管理端则对中间报警进行收集和分析，并基于本文设计的算法对中间报警进行聚合。

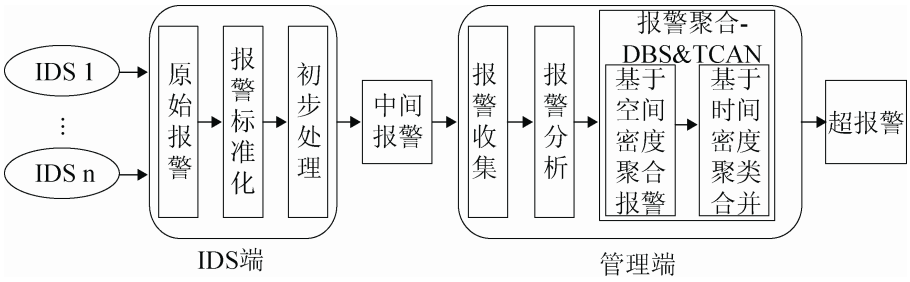


图 1 报警聚合模型

报警预处理是在 IDS 端设置一个缓冲区，选择合适的报警聚合参数和滞留时间对标准化报警进行聚合，能将某一 IDS 针对同一安全事件产生的大量相同或相似的报警聚合成中间报警，可以初步减少一部分重复报警，降低网络负载，且有利于管理端进一步融合分析异构 IDS 产生的报警信息，例如监控某一时间段的攻击，监控特定网络地址的攻击事件等。

DBS&TCAN 算法首先基于空间密度聚合局部报警信息，再基于时间密度对局部聚合结果进行合并，能够将异构 IDS 产生的报警相互补充、印证，进一步减少重复报警，降低误报率，且有利于理解攻击的实质。

1.1 报警标准化

依据 IDMEF(Intrusion Detection Message Exchange Format，入侵检测信息交换格式)定义报警信息格式，即，报警=(事件号，攻击类型，优先

第 28 卷第 6 期 系统仿真学报 Vol. 28 No. 6
2016 年 6 月 Journal of System Simulation Jun., 2016

级，协议，时间戳，检测器编号，源地址，目的地址，源端口，目的端口，信息)，形式化为 Alert =(ID , Type , Priority , Protocal , TimeStamp , SenserID , SrcIP , DstIP , SrcPort , DstPort , Msg)。

参考 IDMEF 模型结构，将报警信息封装成包的形式，如图 2 所示。

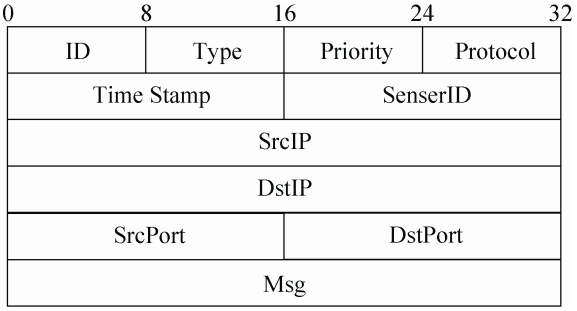


图 2 报警信息格式

1.2 预处理

选择 SrcIP 和 Type 参数作为报警聚合参数，将同属于一个攻击阶段的报警聚合在一起，且聚合粒度有利于报警的进一步处理。通过动态调整报警滞留时间，对短时间内产生重复报警越多的报警类型，使其滞留时间越长，将更多相同的报警聚合在一起，减轻网络负载；反之，对短时间内产生重复报警越少的报警类型，使其滞留时间越短，有利于报警的及时传送。同时，设置 $t_{\max}$ 和 $t_{\min}$ 两个时间阈值，保证报警滞留时间不会因过大或过小而影响报警聚合的实时性和效果。

符号对照如表 1。

表 1 符号对照表

参数	说明
a_i	报警
i, j	攻击类型
t	报警滞留时间
B	缓冲区
$t_{\text{threshold}}$	报警滞留时间阈值
$t_{\max}$	最大滞留时间
$t_{\min}$	最小滞留时间
Δt	滞留时间修正量

Algorithm1：报警预处理
输入：标准化的原始报警
输出：中间报警
1 Begin
2 while new alert a_i is received do
3 if $B=\varnothing$
4 $B=\{a_i\}$
5 else
6 if $\exists a_j \in B$ and $a_{i-SrcIP}=a_{j-SrcIP}$
7 if $i=j$ then
8 $merge(a_i,a_j)$
9 $num_{a_i}=num_{a_i}+1$
10 if $t_{i\text{-}threshold} \leqslant t_{\max}$ then
11 $t_{i\text{-}threshold}=t_{i\text{-}threshold}+\Delta t$
12 else $t_{i\text{-}threshold}=t_{\max}$
13 else
14 $B=B\setminus a_j$ and $B=B\cup\{a_i\}$
15 if $t_{j\text{-}threshold} \geqslant t_{\min}$ then
16 $t_{j\text{-}threshold}=t_{j\text{-}threshold}-\Delta t$
17 else $t_{j\text{-}threshold}=t_{\min}$
18 else $B=B\cup\{a_i\}$
19 every $t_{\min}$，retrieval and send alert in buffer whose $t \geqslant t_{threshold}$
20 end

注：$t_{\max}$、$t_{\min}$、$t_{threshold}$ 和修正量 Δt 设有缺省值。

1.3 报警分析

报警分析主要是通过距离度量来计算报警的相似性，当距离小于一定阈值时，对报警进行聚合。DBS&TCAN 算法涉及空间和时间密度，需要分别计算空间和时间距离。

1.3.1 空间距离计算

2 个报警的空间距离度量方法采用欧几里得距离(Euclidean Distance, ED)，如式(1)所示

$$dist(i,j)=\sqrt{(x_{i1}-x_{j1})^2+(x_{i2}-x_{j2})^2+\cdots+(x_{in}-x_{jn})^2} \quad (1)$$

其中，$i=(x_{i1},\cdots,x_{in})$ 和 $j=(x_{j1},\cdots,x_{jn})$ 是 2 个 n 维对象。为了防止由于输入属性取值范围过大，而范围小的属性对结果的影响被掩盖，可以通过数据标准化，将数据投影到(0,1)区间，再进行运算。

对于报警 Alert =(ID, Type , Priority , Protocal , TimeStamp , SenserID , SrcIP , DstIP , SrcPort , DstPort , Msg)，

(1) ID , Type , Priority , Protocal , SenserID , Port , Msg 等属性，相同则距离为 0，否则为 1；

(2) *IP* 属性距离计算方法如式(2)所示，其中，$sim(ip_i,ip_j)$ 为 ip_i 和 ip_j 中最高位相同的位数。

$$dist(ip_i,ip_j)=1-sim(ip_i,ip_j) \tag{2}$$

(3) TimeStamp 属性距离计算方法如式(3)所示，$T_{\min}$ 和 $T_{\max}$ 为报警的最小和最大时间间隔。

$$dist(t_i,t_j)=\begin{cases}0 & |t_i-t_j|\leqslant T_{\min}\\ 1-\dfrac{T_{\max}-|t_i-t_j|}{T_{\max}-T_{\min}} & T_{\min}<|t_i-t_j|<T_{\max}\\ 1 & \text{otherwise}\end{cases} \tag{3}$$

由于报警各属性的重要度不同，计算空间距离时，结合不同属性的权重 $w_1,w_2,\cdots,w_n$ ($\sum_{p=1}^{n}w_p=1$)，对欧几里得公式进行变换。

$$dist(i,j)=\sqrt{w_1\cdot(x_{i1}-x_{j1})^2+w_2\cdot(x_{i2}-x_{j2})^2+\cdots+w_n\cdot(x_{in}-x_{jn})^2} \tag{4}$$

用一个 $n\times n$ 矩阵存储 n 个报警之间的距离，以便于报警距离计算，即：

$$dist_{n\times n}=\begin{vmatrix}0 & & & \\ d_{21} & 0 & & \\ d_{31} & d_{32} & 0 & \\ \cdots & \cdots & \cdots & \\ d_{n1} & d_{n2} & \cdots & 0\end{vmatrix} \tag{5}$$

$d_{ij}\in[0,1]$，且越接近 0，两个报警越相似。

1.3.2 时间距离计算

两个报警时间距离 $dist_t=|t_A-t_B|$，其中 t_A 和 t_B 分别为报警 *A* 、*B* 的时间戳，即时间距离为两个报警时间戳的差值。

2 DBS&TCAN 算法

参考 DBSCAN 算法的相关定义，对本文报警聚合算法(DBS&TCAN)作如下定义。

定义 1：对象 $p\in D$，以 p 为中心，半径为 *Esp* 内的区域称为 p 的 *Esp* 邻域，记为 $P_{Eps}(p)=\{q\in D\,|\,dist(p,q)\leqslant Eps\}$，其中 *D* 为数据集合，$dist(p,q)$ 表示对象 p 和 q 的距离。

定义 2：给定对象 p，半径为 *Esp* 的邻域内的样本点数大于等于 *MinPts*，则称 p 为核心对象。

定义 3：对于数据集合 *D*，若 $q\in P_{Eps}(p)$，且 $|P_{Eps}(p)|\geqslant Minpts$，那么对象 q 从对象 p 直接密度可达。

定义 4：对象 q 从对象 p 直接密度可达，且 q 为非核心对象，则 q 为边界对象。

定义 5：不属于任何类的对象称为噪声。

2.1 算法描述

设系统实时处理报警的时间间隔为 T_0，报警聚合半径为 ε_{1i} 和 ε_2，其中 ε_{1i} ($i=1,2,\cdots$) 为基于密度的空间聚类半径，为不同时间间隔内报警的聚合半径；ε_2 为基于密度的时间聚类半径，取 $\varepsilon_2=T_0$。假设 T_0 时间内有 *m* 个报警到达，记为 $a_1,a_2,\cdots,a_m$。

算法按报警最近到达最先处理的原则，从队尾 a_m 开始处理。算法分为两部分，首先选择局部聚合参数 ε_{1i} 对 T_0 内的报警进行聚合；再设定聚合时间跨度 $\Delta T=N\cdot T_0$，对满足条件(Algorithm2 第 10 步)的局部聚合结果进行合并，直到没有直接密度可达的报警可以聚合或超出时间范围。

$Aggr(a_1,\cdots,a_m)=c_1\cdot f_1(a_1,\cdots,a_m)+c_2\cdot f_2(a_1,\cdots,a_m)+\cdots+c_n\cdot f_n(a_1,\cdots,a_m)$，其中 $c_1,c_2,\cdots,c_n$ 为常数，$f_1,f_2,\cdots,f_n$ 为属性聚合函数，包括时间戳、优先级等属性的更新与合并，采用加权平均的方法更新聚合报警的参数。

2.2 局部参数选择-ε_{1i}

对 T_0 内的报警构建 R^*-树(降低搜索的时间复杂度)和 $dist_{n\times n}$ 矩阵。对 $dist_{n\times n}$ 矩阵的列进行排序并转置得到矩阵 $dist_{n\times n}^{tran}$，其列向量代表对象到最近的第 $k-1$ ($k=1,2,\ldots,n$) 个对象的距离集合。其中第一列表示每个对象到它自身的距离，即第一列全零。删除第一列得到矩阵 $dist_{n\times(n-1)}^{tran}$，其第 k ($k=1,2,\ldots,n$) 列是报警集合中所有对象的 *k*-最邻近距离集合 $Dist_k$ (排序后即为 *k-dist* 图)。

第 28 卷第 6 期 系统仿真学报 Vol. 28 No. 6
2016 年 6 月 Journal of System Simulation Jun., 2016

Algorithm2：DBS&TCAN
输入：中间报警
输出：超报警
1 **Begin**
2 **FOR** $k=m$ to 1
3 **if** $P_{\varepsilon_{1i}}(a_k)=\varnothing$ **then**
4 $A_k=\{a_k\}$
5 **else**
6 $A_k=P_{\varepsilon_{1i}}(a_k)=\{a_x,\cdots,a_k,\cdots,a_y\}$
7 $A_{k-aggr}=Aggr(A_k)$
8 **FOR** $k=m$ to 1 **do**
9 **int** $\Delta T=N\cdot T_0$
10 **if** $\left|t_{A_k}-t_{A'}\right|\leqslant\varepsilon_2$ and $\exists q\in A_k, p\in A'$ $dist(p,q)\leqslant\min(\varepsilon_{1i},\varepsilon_1')$ **then**
11 $Aggr(A_k,A')$
12 $N=N-1$
13 $A_k=A'$
14 **Until** $A'=\varnothing$ or $N=0$
15 **end**

取 $k=4$ (一般 $Minpts=4$，而 ε_{1i} 由 $k=Minpts$ 的 *k-dist* 图确定)，对于 $Dist_4$ 的概率分布情况，选其峰值点所对应的 *k*-最近邻距离值(横坐标刻度)为 ε_{1i}。

对 $Dist_k$ 的概率分布情况进行拟合，概率公式为 $P(x)=\sqrt{\frac{\lambda_k}{2\pi x^3}}\exp[-\lambda_k(x-\mu_k)^2/(2x\mu_k{}^2)]$ (高斯分布)，其中 λ_k 和 μ_k 可以用 MLE(极大似然估计)获得。为了获得峰值点，解微分方程 $dP(x)/dx=0$ 得到一个正数解，即：

$$\varepsilon_{1i}=\frac{\mu_k\sqrt{9\mu_k{}^2+4\lambda_k{}^2}-3\mu_k{}^2}{2\lambda} \tag{6}$$

2.3 聚类的合并

对 T_0 时间内报警信息进行聚合的过程中，需要记录边界对象与噪声对象信息，因为它们可能是全局中某个聚类的边界对象或某一个被分割的小聚类中的对象。报警聚类的合并可分为以下几种情况：

(1) 对不同时间窗口的报警分别聚合，可能使得本属于同一类的报警类被划分到相邻两个区域，应对这两个类进行合并。对两个类进行合并，当且仅当(Algorithm2 第 10 步)：

$\left|t_{A_k}-t_{A'}\right|\leqslant\varepsilon_2$ and $\exists q\in A_k, p\in A'$，$dist(p,q)\leqslant\min(\varepsilon_{1i},\varepsilon_1')$

(2) 处于 T_0 时间窗口边界附近的噪声可能是其他类的边界对象，应将这些噪声归并到某个类中。噪声对象 a_j 与某个类进行合并，当且仅当：

$\left|t_{A_k}-t_{a_j}\right|\leqslant\varepsilon_2$ and $\exists q\in A_k$，$dist(p,a_j)\leqslant\varepsilon_{1i}$

(3) 在不同的 T_0 时间窗口，一些小的报警聚类可能被划分到不同窗口中去。由于每个窗口的报警数量过小，在各分区的聚合过程中被标记为噪声对象，应将这些噪声聚合成新类。对于噪声对象的处理办法如下：

设两个相邻时间窗口内的噪声对象集合为 S，两个时间窗口报警聚合的空间半径为 $\varepsilon_{1i},\varepsilon_1'$。若 $\exists p_0\in S, q_i\in S(i=1,2,\cdots,n,n\geqslant Minpts)$，且 $p_0\neq q_i$，满足 $\Delta t\leqslant\varepsilon_2$，$dist(p,q_i)\leqslant\min(\varepsilon_{1i},\varepsilon_1')$，则以 p_0 为核，$\{q_1,q_2,\cdots,q_n\}$ 为边界对象形成一个新的聚类。

3 实验及结果分析

为了验证算法有效性，实验基于异构入侵检测系统，并采用数据集测试的方法对算法性能进行了测试与分析。

3.1 实验方案

3.1.1 实验平台

实验采用多引擎协作入侵检测，包括 Bro[12]，Lerad[13]，Snort[14]三种 IDS，如表 2 所示。

表 2 入侵检测系统

检测系统	分类	描述
Bro	误用	核心是事件的生成和处理机制，预定义一系列的基本事件，同时对这些事件注册基本的事件处理函数。
Lerad	异常	通过学习算法从训练数据集中学习正常的规则集，并以此作为实际检测的规则集。
Snort	误用	采用基于规则的工作方式，可以检测 1 000 多种不同的入侵行为和探测活动，但误报率较高。

3.1.2 实验数据

实验数据集采用美国国防部高级计划研究署

提供的 DARPA 99 数据集，包括 Probe，DoS，R2L，U2R 和 Data 等 5 大类 58 种典型攻击方式，是目前最为全面的攻击测试数据集。如表 2 所示，该评测数据集给出了五周模拟数据，其中前两周数据用于训练，而后三周数据则用于评测。第一、三周为不包含任何攻击的正常数据，第二周插入属于 18 种类型的 43 次攻击实例，第四、五周包含属于 58 种类型的 201 次攻击实例。

实验采用 DARPA 99 第一周数据作为 Leard 的训练集，训练过程离线进行。测试数据集采用第四周每天的 Tcpdump 类型数据，包括超过 20 种类型的 104 次攻击实例。通过数据重放工具(如 Tcpreplay、Netpoke 等)对每天的测试数据进行重放，由 3 种入侵检测系统产生原始报警信息，并实时处理。

3.2 实验结果

为验证实验结果，定义精简率=聚合减少报警数/总报警数。实验结果数据如表 3 所示，包括预处理和报警聚合后剩余报警数量以及相应的精简率。

表 3 聚合结果分析

数据 (DARPA 99_4)	原始报警	预处理	预处理精简率%	报警聚合	报警聚合精简率%	总精简率%
周一	Bro：2 451	889	63.7	462	86.8	94.7
	Lerad：3 634	1 617	55.5			
	Snort：2 606	1 005	61.4			
周二	Bro：2 212	861	61.1	383	89.2	95.7
	Lerad：4 022	1 654	58.9			
	Snort：2 661	1 058	60.2			
周三	Bro：2 369	977	58.8	324	90.7	96.2
	Lerad：3 587	1 527	57.4			
	Snort：2 534	963	62.0			
周四	Bro：2 524	903	64.2	393	88.8	95.8
	Lerad：3 864	1 497	61.2			
	Snort：2 763	972	64.8			
周五	Bro：2 173	762	64.9	279	90.9	96.5
	Lerad：3 438	1 422	58.6			
	Snort：2 264	873	61.4			

由表 3 可知，对单个 IDS 的报警进行预处理，精简率在[55.5, 64.9]之间；采用 DBS&TCAN 算法对不同 IDS 的报警进行聚合，精简率在[86.8, 90.9]之间。总精简率在[94.7, 96.5]之间，大大减少了报警数量。

扫描攻击和 Dos 攻击等易在短时间内诱发大量报警，本文所设计的算法对扫描攻击和 Dos 攻击的检测更有效。图 3 给出了部分扫描攻击的检测结果，包括对 FTP 端口和 HTTP 端口的水平扫描和针对两个主机的垂直扫描，可以看出算法对扫描攻击产生报警聚合效果较好，为进一步分析攻击者意图奠定了基础。

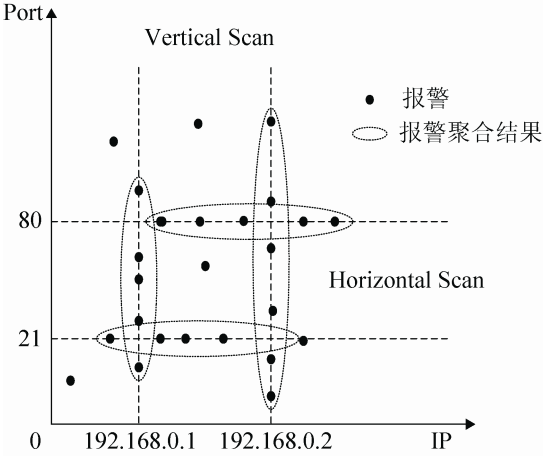


图 3 扫描攻击聚合结果

第 28 卷第 6 期 系统仿真学报 Vol. 28 No. 6
2016 年 6 月 Journal of System Simulation Jun., 2016

表 4 给出了不同 IDS 的检测率和误报率，以及报警聚合分析处理后的数据对比，实验证明本文设计的报警聚合算法能有效提高系统检测率和降低系统误报率。

表 4 检测率与误报率对比

性能对比%	Bro	Lerad	Snort	本文
检测率	35.7	56.2	42.1	71.4
误报率	31.4	42.6	27.3	1.1

结合表 3 和表 4 进行分析，可以发现，本文在保证大大降低报警数量的同时也提高了报警质量，即滤除大量重复报警和误报警的同时，提高了系统检测率，保证了系统的有效性和可用性。

图 4 给出了 ROC 曲线图，并标出不同 IDS 工作的状态点，基线以 Snort 系统为基准，本文方案在检测率和误报率上均有明显改善。

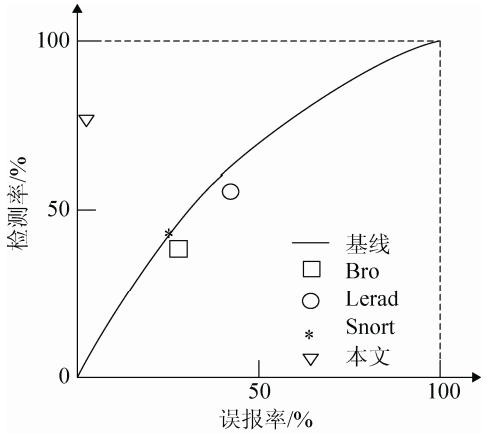


图 4 ROC 图

3.3 性能分析

采用 DARPA 99 第四周数据作为测试数据集，分别对文献[9-11]中的方案进行测试，并与本文方案进行对比，如表 5 所示。其中，m 为 T_0 时间内的报警数据增量，n 为内存中已有的报警数据量，γ 为合并局部聚类的操作数。

表 5 相关工作比较

	文献[9]	文献[10]	文献[11]	本文
计算复杂度	$(m+n)\lg(m+n)$	$(m+n)\lg(m+n)$	$(m\cdot n)$	$m\lg m+\gamma$
精简率%	[65.1, 73.4]	[83.7, 87.1]	[82.3, 89.0]	[94.7, 96.5]
误报率%	6.3	1.9	2.4	1.1
检测率%	52.1	70.3	66.2	71.4
实时性	差	差	较好	较好

由表 5 可知，本文方案在具有较高精简率和较低误报率的同时，也具有较高的检测率，并在实时性方面体现出优势。

4 结论

针对分布式多引擎入侵检测系统在实际应用中存在大量重复报警和高误报率问题，本文提出了一种基于空间和时间密度的抗噪声聚合算法(DBS&TCAN)。算法采用并行处理思想，对不同时间窗口内的报警并行使用 DBSCAN 算法，有效提高了数据处理速度，降低 I/O 和内存开销，而且改善了由报警信息密度不均匀导致聚合效果差的不足；在合并局部报警聚类的过程中，引入时间密度，充分考虑噪声对象，对一定时间跨度内的聚类进行合并。实验结果表明，算法能够减少大量重复报警和误报警，提高了系统的检测率和时效性。

参考文献:

[1] 卿斯汉, 蒋建春, 马恒太, 等. 入侵检测技术研究综述 [J]. 通信学报, 2004, 25(7): 19-29.

[2] 穆成坡, 黄厚宽, 田盛丰. 入侵检测系统报警信息聚合与关联技术研究综述 [J]. 计算机研究与发展, 2006, 43(1): 1-8.

[3] T Kanungo, N S Netanyahu, A Y Wuan. An Efficient k-Means Clustering Algorithm: Analysis and Implementation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence(S0162-8828), 2002, 24(7): 881-892.

[4] Hadi Bahrbegi, Ahmad Habibizad Navin, Amir Azimi Alasti Ahrabi. A New System to Evaluate GA-based Clustering Algorithms in Intrusion Detection Alert

Management [C]// 2010 Second World Congress on Nature and Biologically Inspired Computing. Kitakyushu (Japan): Institute of Electrical and Electronics Engineers (IEEE), 2010: 115-120.

[5] A Hofmann, D Fisch, B Sick. Identifying Attack Instances by Alert Clustering [C]// Proc. IEEE Three-Rivers Workshop Soft Computing in Industrial Applications. USA: IEEE, 2007: 25-31.

[6] G C Tjhai, S M Furnell, M Papadaki, et al. Preliminary two-stage alarm correlation and filtering system using SOM neural network and K-means algorithm [J]. Computers & Security(S0167-4048), 2011, 29(6): 712-723.

[7] Emmanuel Hooper. An Intelligent Intrusion Detection and Response System Using Hybrid Ward Hierarchical Clustering Analysis [C]// Multimedia and Ubiquitous Engineering. Crete(Greece): Institute of Electrical and Electronics Engineers (IEEE), 2010: 1187-1192.

[8] C Tsai, C C Yen. Unsupervised Anomaly Detection Using HDG-Clustering Algorithm [C]// Neural Information Processing, Germany: Springer Berlin (Heidelberg), 2009: 356-365.

[9] M Ester, H P Kriegal, J Sander, et al. A density-based algorithm for discoverying clusters in large spatial databases with noise [C]// KDD96 Proceedings of 2nd International Conference on Knowledge Discovery and Data Mining. Portland(Oregon): AAAI Press, 1996, 96(34): 226-231.

[10] Tran Manh Thang, Juntae Kim. The Anomaly Detection by Using DBSCAN Clustering with Multiple Parameters [C]// Information Science and Applications (ICISA), 2011 International Conference (IEEE). USA: IEEE, 2011: 1-5.

[11] Alexander Hofmann, Bernhard Sick. Online Intrusion Alert Aggregation with Generative Data Stream Modeling [J]. IEEE Transactions on Dependable and Secure Computing(S1545-5971), 2011, 8(2): 282-294.

[12] Xfocus Team. Bro: 一个开放源码的高级 NIDS 系统. [EB/OL]. (2003-10-12) [2015-04-29]. http://www.xfocus.net/articles/200310/624.html.

[13] M V Mahoney, P K Chan. Learning rules for anomaly detection of hostile network traffic [C]// Proc. 3rd IEEE Int'l Conf. Data Mining Los Alamitos, CA, USA. USA: IEEE Computer Society Press, 2003: 601-604.

[14] Caswell B, Roesch M. Snort: The open source network intrusion detection system [EB/OL]. (2009-04-21) [2015-04-29]. http://www.snort.org/

[15] 穆成坡, 黄厚宽, 田盛丰. 入侵报警管理与入侵响应系统&中的自适应报警聚合 [J]. 计算机科学, 2007, 34(12): 73-77.

[16] 胥小波, 蒋琴琴, 郑康锋, 等. 基于混沌粒子群的告警聚类算法 [J]. 通信学报, 2013, 34(3): 105-110.

(上接第 1335 页)

4 结论

本文提出了以四轴转台作为基础的方位可连续旋转的垂直发射飞行器姿态运动的仿真方法，并设计了四轴协调运行控制规律，综合了立式三轴转台和卧式三轴转台的优势，避免了两者的缺陷，为方位可连续旋转的垂直发射飞行器姿态运动的实时模拟提供了一种连续、实时、高效的仿真方法。

参考文献:

[1] 常卫伟. 舰载导弹垂直发射系统巡礼 [J]. 舰载武器, 2004 (1): 65-68.

[2] 郑宏建. 舰载导弹垂直发射与安全性分析 [J]. 飞航导弹, 2009 (2): 16-19.

[3] 李跃军, 阎超. 飞行器姿态角解算的全角度双欧法 [J]. 北京航空航天大学学报, 2007, 33(5): 505-508.

[4] 王亚东, 袁绪龙, 张宇文. 双欧控制法在运载器水弹道中的应用 [J]. 鱼雷技术, 2013 (6): 401-405.

[5] 马杰, 姚郁. 新型仿真转台设计方案研究 [J]. 系统仿真学报, 2009, 21(增2): 129-132.

[6] 孟凡伟. 三轴飞行仿真转台伺服控制系统设计研究 [D]. 哈尔滨: 哈尔滨工业大学, 2001: 1-56.

[7] 吴云洁. 飞行仿真转台的完全跟踪控制 [J]. 控制理论与应用, 2011, 28(3): 414-420.