一种基于递归图的网络时间隐蔽信道检测方法

刘 标[1] 兰少华[1] 张 晶[1] 刘光杰[2]
(南京理工大学计算机科学与工程学院 南京 210094)[1] (南京理工大学自动化学院 南京 210094)[2]

摘 要 网络时间隐蔽信道的检测是网络隐蔽信道研究中的热点和难点。熵检测是目前最有效的检测方法,可有效检测多种网络时间隐蔽信道。但随后提出的Liquid隐蔽信道采用熵补偿的方法有效地躲避了熵检测。提出了一种基于递归图的检测算法,其可以检测出包括Liquid在内的多种网络时间隐蔽信道。

关键词 时间隐蔽信道,递归图,网络安全
中图法分类号 TP393.08 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2015.2.024

Approach Based on Recurrence Plot to Detect Covert Timing Channels

LIU Biao[1] LAN Shao-hua[1] ZHANG Jing[1] LIU Guang-jie[2]
(School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing 210094, China)[1]
(School of Automation, Nanjing University of Science and Technology, Nanjing 210094, China)[2]

Abstract The detection of covert timing channel is the focus of the research of covert channel, and it is very difficult. Entropy-based approach is the most effective detection approach. It can detect almost all the covert timing channels. However, Liquid is proposed soon after. It can effectively evade the entropy-based detection by smoothing the entropy. In this paper, a detection approach based on recurrence plot was introduced, and the detection approach can detect various covert timing channels including Liquid.

Keywords Covert timing channel, Recurrence plot, Network secure

1 引言

Lampson Butler于1973年第一次给出了隐蔽信道的定义,定义指出隐蔽信道是指违背设计者初衷而被设计用来通信的信道[1]。文献[2]第一次将隐蔽信道的概念扩展到了网络。网络隐蔽信道在APT攻击中是主要的通信手段,严重威胁网络安全。在网络中,根据传输载体的不同,可将隐蔽信道分为网络存储隐蔽信道和网络时间隐蔽信道[3-10]。网络时间隐蔽信道的检测一直是该领域的难点,近年来,国内外的研究者提出了一些针对网络时间隐蔽信道的检测方法[3,11-13],然而这些检测方法中有些只能检测特定的隐蔽信道,有些检测方法虽能检测多种隐蔽信道,但对网络环境高度敏感。Gianvecchio等人提出的熵检测方法是公认的最有效的检测方法,可以检测出多种隐蔽信道,且具有较高的检测率。然而,Robert等人设计了一种可以躲避熵检测的隐蔽信道Liquid[8],其通过熵补偿的办法有效应对了熵检测。

本文提出一种基于递归图的网络时间隐蔽信道检测方法。递归图是一种有效的非线性时间序列定性分析工具,它可以在二维图形上反映非线性动力系统的高维(维度大于3)相空间轨道特性,进而揭示非线性动力系统的动力学演化规律。递归图分析技术自提出以来,广泛应用于物理现象和生理机理等采样信号序列的分析上。网络数据包间时间间隔序列IPDS(Inter Packet Delays Series)由于受网络应用程序、网络环境、服务器响应等多方面复杂因素的影响,本身具有较强的非线性和非评估特性。可以利用递归图方法刻画IPDS的动力学特性,通过检查动力学特性是否在隐蔽信道形成过程中受到破坏,来判断是否存在隐蔽信道。实验结果表明,该方法可以有效地检测多种隐蔽信道,包括抗熵检测的Liquid隐蔽信道。

2 相关工作

早期,对于网络时间隐蔽信道的研究主要集中在信道的干扰和消除上,通过给信道增加噪声来降低信道容量,从而限制网络时间隐蔽信道的容量。在后来的研究中,研究人员逐渐倾向于研究网络时间隐蔽信道的构造和检测,接下来也将详述已有的网络时间隐蔽信道的构造和检测方法。

2.1 网络时间隐蔽信道构造

网络时间隐蔽信道可分为主动式隐蔽信道和被动式隐蔽信道。主动式隐蔽信道会主动产生用于构造隐蔽信道的网络数据流,而被动式隐蔽信道则通过操控已经存在的数据流中

到稿日期:2014-07-10 返修日期:2014-08-30 本文受国家自然科学基金(61170250,61103201),中央高校基本业务费专项资金(30920140121006)资助。

刘 标(1989-),男,硕士生,主要研究方向为计算机网络和安全,E-mail:929037165@qq.com;兰少华(1958-),男,博士,教授,主要研究方向为计算机网络及应用、网络安全、分布式人工智能;张 晶(1988-),女,硕士生,主要研究方向为计算机网络和安全;刘光杰(1980-),男,博士,副研究员,主要研究方向为网络与信息安全。

的时间间隔来进行隐蔽信息的传输。一般来讲，被动隐蔽信道相对较难检测。

2.1.1 主动式隐蔽信道

Cabuk 等人在文献[3]中提出了第一种基于 IP 的网络时间隐蔽信道，称作 IPCTC（IP Covert Timing Channel）。IPCTC 通过判断一个时间间隔内是否发送数据包来表达数据 1 和 0；为了克服 IPCTC 产生的 IPDS 与正常网络数据流产生的 IPDS 分布差异较大的问题，Cabuk 在文献[4]中提出了另一种主动式网络时间隐蔽信道 TRCTC（Time-Replay Covert Timing Channel），因为 TRCTC 使用了正常网络数据流的 IPDS 作为输入样本，所以这种信道编码机制产生的 IPDS 的分布和正常网络数据流的 IPDS 分布很接近，抗检测性能较好；Gianvecchio 等人在文献[5]中提出了一种基于模型的主动式网络时间隐蔽信道 MBCTC（Model-Based Covert Timing Channel），MBCTC 对正常网络数据流时间间隔进行分布拟合，根据拟合的分布产生时间间隔，因此 MBCTC 产生的 IPDS 的分布几乎和正常网络数据流的 IPDS 分布一致，抗检测性能极高；Liu 等人在文献[6]中提出了一种分布匹配的隐蔽信道 DMCTC，这种隐蔽信道可以有效抵抗基于统计方法的检测。

2.1.2 被动式隐蔽信道

Shah 等人在文献[7]中提出了一种使用硬件实现的网络时间隐蔽信道 JitterBug，JitterBug 通过一个硬件模块对键盘的触发事件稍作延时，从而调制交互式程序（如 SSH 等）数据包的发送时间，达到信道编码的目的，JitterBug 相当于给正常网络数据流的 IPD 增加了一个较小的抖动，使得 JitterBug 产生的 IPDS 与正常的 IPDS 非常相似，抗检测性能较高；Robert等人在 JitterBug 的基础上提出一种抗熵检测的隐蔽信道构造算法 Liquid[8]，Liquid 在 JitterBug 的基础上增加了一个平滑熵变化的环节，从而使熵检测算法失效。

2.1.3 其他网络时间隐蔽信道

Luo 等人在文献[9]中提出了一种基于多条数据流的隐蔽信道 Cloak，随后 Luo 等人又在文献[10]中提出一种建立在 TCP 层的网络时间隐蔽信道 TCPScript，这种方法通过 ACK 包来确保隐蔽信息的正确性。

2.2 网络时间隐蔽信道检测

网络时间隐蔽信道的检测方法可分为专用检测方法和通用检测方法，专用检测方法只能检测出某种特定的隐蔽信道，而通用的检测算法可以检测出多种隐蔽信道。

2.2.1 专用检测方法

Cabuk 在文献[3]中针对 IPCTC 提出了两种基于规律度的检测方法。第一种方法根据时间间隔的标准差的变化程度来判断是否存在隐蔽信道，该方法认为正常的 IPDS 的标准差变化较大，而 IPCTC 的 IPDS 在编码方法不变的情况，下标准差变化不大。Cabuk 提出的另一种基于规律度的检测方法称作 ε-similarity，该方法通过衡量相似时间间隔的比例来判断是否存在隐蔽信道，ε-similarity 检测方法认为 IPCTC 会产生相似时间间隔的聚集效应。Luo 在文献[9]中提出了针对 Cloak 的检测方法，该方法通过衡量数据包和应答包之间的时间间隔来判断是否存在隐蔽信道。

2.2.2 通用检测方法

Peng 等人在文献[11]中提出了基于 Kolmogorov-Smirnov 试验的检测方法，该方法首先计算测试样本和训练样本的经验分布函数的距离，如果距离较大，则说明测试样本为隐蔽信道，反之则不是隐蔽信道。Berk 等人在文献[12]中讨论了一种基于区间统计的检测方法，通过计算包间时间间隔均值所在区间的时间间隔数目 C_μ 和峰值区间的时间间隔数目 $C_{\max}$ 的差距 $1-C_\mu/C_{\max}$ 来作为判断隐蔽信道的标准。Gianvecchio 等人在文献[13]中提出了目前公认最有效的检测方法：熵检测，作者认为时间隐蔽信道的检测主要分为两类：基于形状的检测和基于规律的检测，IPDS 的形状可以通过一阶统计来描述，而时间间隔序列的规律可以通过高阶统计来描述。Gianvecchio 等人认为构造隐蔽信道必然会影响正常 IPDS 的熵，基于此，他们提出了利用熵率和修正条件熵来检测隐蔽信道，熵率可以描述时间间隔序列的形状，而修正条件熵可以描述时间间隔序列的规律，通过这两种熵可以有效检测出绝大多数时间隐蔽信道。

3 基于递归图的检测方法

本文提出的基于递归图的检测方法旨在检测出多种隐蔽信道，包括抗熵检测的 Liquid 隐蔽信道。在这一部分中，我们将首先介绍递归图的概念，然后详述基于递归图检测方法的设计。

3.1 递归图

相空间重构是分析非线性系统的重要步骤，Takens 等人在文献[14]中给出了一种只需一维时间序列就可以重构非线性系统的方法，该方法按式(1)将一维时间序列 $x_i(i=1,2,3,\cdots,N)$ 重构为 m 维相空间，向量 X_i 表示第 i 个相点，m 为嵌入维数，m 的值可通过错误最邻近法（FNN）或奇异值分析法等确定，τ 为延迟时间，τ 的值可通过互信息法或自相关函数法等确定。

$$X=\begin{Bmatrix}X_1\\X_2\\\vdots\\X_M\end{Bmatrix}=\begin{Bmatrix}x_1 & x_{1+\tau} & \cdots & x_{1+(m-1)\tau}\\x_2 & x_{2+\tau} & \cdots & x_{2+(m-1)\tau}\\\vdots & \vdots & \vdots & \vdots\\x_{N-(m-1)\tau} & x_{N-(m-2)\tau} & \cdots & x_N\end{Bmatrix} \tag{1}$$

递归图是基于相空间重构的非线性系统定性分析方法，递归图的概念由 Eckmann 等人在文献[15]中提出，它能够在二维空间上直观地反映非线性系统的高维相空间运动规律。递归图由二维方阵中的白点和黑点组成，二维方阵中的白点表示两个状态点远离，黑点表示两个状态点逼近。递归图的数学表达式为：

$$R_{i,j}=\Theta(\varepsilon-\|X_i-X_j\|),i,j=1,2,\cdots,M \tag{2}$$

其中，ε 为设定距离阈值，X_i 和 X_j 分别表示第 i 个和第 j 个相点，$\|\cdot\|$ 表示范数（如 Euclidean 距离），$\Theta(\cdot)$ 是 Heaviside 函数，M 为相点的个数。当 X_i 和 X_j 的距离小于 ε 时，$R_{i,j}$ 为 1，递归图中 (i,j) 位置上为黑点，否则 (i,j) 位置上为白点。

递归图的结构受距离阈值 ε 的影响较大，ε 的选取常常需要领域专家根据经验进行确定。为了克服这个缺点，可使用排序递归图（ORP），排序递归图根据相点的排序模式来确定相点间的逼近或远离关系，定义如式(3)所示[16]：

$$R_{i,j}=\begin{cases}0, & \pi_i\neq\pi_j\\1, & \pi_i=\pi_j\end{cases},i,j=1,2,\cdots,N \tag{3}$$

其中，π_i 和 π_j 分别表示第 i 个和第 j 个相点的排序模式，m

维的相点可能的排序模式有 $m!$ 种。在二维图中分别用白点和黑点表示 $R_{i,j}$ 值为 0 和 1 的情况，即可绘出排序递归图。

图 1、图 2 分别为正常 HTTP 数据流和 IPCTC 的 IPDS 的排序递归图。从图 1 中可以看出，大部分的递归点是沿着平行于对角线的方向分布的，这说明 HTTP 数据流的 IPDS 确定性比较强。而图 2 中的递归点比较散乱，说明 IPCTC 的 IPDS 确定性较差。

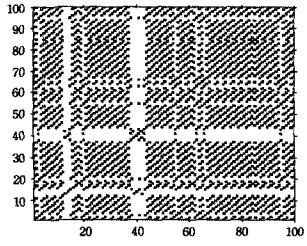

图 1　正常 HTTP 的排序递归图

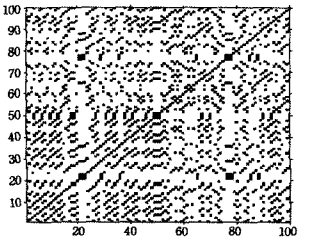

图 2　IPCTC 的排序递归图

3.2　定量递归分析

递归图只能用来对系统进行定性分析，这不适合于在线应用。Zbilut 和 Webber 等人在文献[16]中首次提出了递归图的量化分析方法，随后研究人员针对递归图的特性陆续提出了不同的量化分析方法，主要的量化分析特征量有递归律、确定率、发散度、熵、层流率、第一及第二递归次数等。针对网络时间隐蔽信道的检测，我们选择的量化分析特征量为确定率。

确定率(DET)的定义为：平行于对角线线段的递归点的个数与总的递归点的比值。计算方法如式(4)所示：

$$DET=\frac{\sum_{l=l_{\min}}^{N} lP(l)}{\sum_{l=1}^{N} lP(l)} \tag{4}$$

其中，$l_{\min}$ 为对角线线段长度，一般取 $l_{\min}=2$，$P(l)$ 为长度为 l 的对角线的频率。确定性特征量可以用来量化系统的确定性。

3.3　检测算法

本文根据排序递归图的确定率(DET)这个特征量来区分正常数据流和隐蔽信道数据流，具体检测算法实现步骤如下：

(1)确定检测窗口 w，计算大量正常样本在窗口 w 下的 DET 值。

(2)选取检测阈值，确定正常数据流的 DET 值域，使得正常样本的误检率在 5%以内。

(3)计算测试数据在窗口 w 下的 DET 值。

(4)如果计算的 DET 在正常阈值值域之外，则为隐蔽信道，否则为正常信道。

由于正常数据流复杂多样，不同的数据流得出的检测阈值有所不同，应区别对待。在实际应用中，应分别计算不同协议的检测阈值，如分别根据 HTTP、FTP、SMTP 和 SSH 协议确定各自的检测阈值。

3.4　实验设计

为了体现本文所提检测算法的通用性，本文将对 5 种隐蔽信道进行检测，分为两种类型，一种为主动式隐蔽信道：IPCTC、TRCTC、MBCTC，另一种为被动式隐蔽信道：Jitter-Bug 和 Liquid。另外，为了体现本文所提出的基于递归图的检测算法的有效性，我们将和熵检测算法进行对比。

3.4.1　数据集

本文选取两种协议类型的数据流作为背景数据流：HTTP 和 SSH。HTTP 协议是目前应用最广泛的协议，防火墙大多都会选择放行 HTTP 协议，因此我们选择 HTTP 协议的数据流作为主动式隐蔽信道的背景数据流，采取上传超大文件的方式来产生大量 HTTP 协议的数据流。由于 Jitter-Bug 和 Liquid 最初的设计是建立在 SSH 协议上的，因此我们选择 SSH 协议作为背景数据流，为了获得大量的 SSH 协议数据流，我们选择网上公开的数据集：Waikato VIII[17]。

3.4.2　相空间重构参数选取

相空间重构的参数为嵌入维数 m 和延迟时间 τ，用于确定参数的正常数据流为上文所提的背景数据流。对于嵌入维数 m，本文采用错误最邻近法来确定，计算得出 HTTP 协议和 SSH 协议的嵌入维数均为 3。对于延迟时间 τ，本文采用互信息法确定，计算得出的 HTTP 和 SSH 的时间延迟均为 1。

3.4.3　隐蔽信道构造

各隐蔽信道将根据参考文献中提供的最高级版本进行构造。对于主动式隐蔽信道，本文采用 200000 个正常 HTTP 数据包作为主动发送的数据包。其中 IPCTC 采用文献[3]建议的轮流使用 40ms、60ms 和 80ms 3 种时间间隔的构造方法；TRCTC 使用 BMC 版本[4]；MBCTC 根据 200000 个正常 HTTP 数据包间时间间隔来拟合分布模型。对于被动式隐蔽信道，本文重放 SSH 数据包，并根据编码算法对数据包进行延时。

4　实验结果分析

首先，我们采集了 5 种隐蔽信道产生的数据包各 200000 个，同时也选取正常 HTTP 数据包和 SSH 数据包各 200000 个，选用熵检测采用的检测窗口 2000 来计算熵检测值和基于递归图的检测值(DET)，熵检测的参数与文献[13]中一致，基于递归图的检测参数选用情况如下：主动式隐蔽信道的检测中对角线线段长度 $l_{\min}$ 选择 2，被动式隐蔽信道的检测中对角线线段长度 $l_{\min}$ 选择 4。这样每种隐蔽信道和正常信道都有 100 组检测值数据。在 4.1 节和 4.2 节会给出这些检测值的分析和对比。接下来，4.3 节会考察检测窗口对检测率的影响。

4.1　主动式隐蔽信道检测结果分析

图 3 给出了 40 组正常 HTTP 数据流和各主动式隐蔽信道的 DET 值。从图 3 中可以看出，正常 HTTP 数据流的确定率(DET)值明显高于各主动式隐蔽信道，说明正常 HTTP 数据流的递归点大部分出现在对角线上，孤立点相对较少，从图 1 中可明显看出此特点，这表明正常 HTTP 数据流的确定性比较强；而各主动式隐蔽信道的确定率(DET)值相对较低，说明递归点多以孤立的形式存在，系统具有较强的随机性。事实上，由于本文选择的正常数据流是基于 HTTP 协议上传附件的形式产生的数据流，正常 IPDS 的模式主要由协议、网络、客户端的处理能力和服务器端的处理能力这 4 个方面决定，这 4 者确定的情况下，正常 IPDS 的确定性就比较强。而主动式隐蔽信道的 IPDS 模式主要由隐蔽信道编码算法决定，否则就会有较高的误码率，由于编码信息的确定性较小，导致了隐蔽信道的 IPDS 的确定性较小。

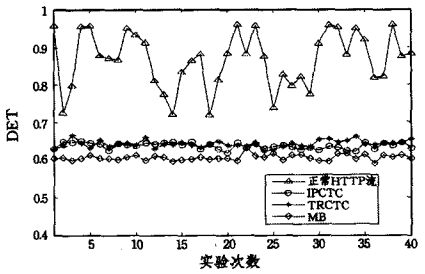

图 3　正常 HTTP 数据流和各主动式隐蔽信道的 DET

表 1 为 100 组实验数据中各主动式隐蔽信道的检测率以及正常 HTTP 信道的虚警率。从表 1 中可以看出，熵检测算法和基于递归图的检测算法都可以有效检测出各主动式隐蔽信道，而且基于递归图的检测算法可在 0%的虚警率的情况下达到 100%的检测率，对主动式隐蔽信道检测效果明显优于熵检测算法。

表 1　主动式时间隐蔽信道检测结果

检测阈值	HTTP 虚警率	IPCTC 检测率	TRCTC 检测率	MBCTC 检测率
EN≤6.6 或 EN>7.5	1%	94%	100%	98%
CCE≥1.9	1%	100%	100%	97%
DET≤0.67	0%	100%	100%	100%

4.2　被动式隐蔽信道检测结果分析

图 4 示出了正常 SSH 流与 JitterBug 和 Liquid 两种被动式隐蔽信道的 40 组 DET 值，从图中我们可以发现正常 SSH 流与两种被动式隐蔽信道的 DET 值很接近，这是因为这两种被动式隐蔽信道只是在正常的时间间隔的基础上增加适当的抖动，在抖动较小的情况下，正常时间间隔序列的动力学特性的变化相对也较小，DET 值就比较接近。与此同时，太小的抖动会导致误码率较高，在保证较小误码率的情况下，正常时间间隔序列的动力学特性必然发生改变，所以图 4 中，虽然正常 SSH 流与两种被动式隐蔽信道的 DET 值很接近，但是正常 SSH 流与两种被动隐蔽信道仍能很好地区分。

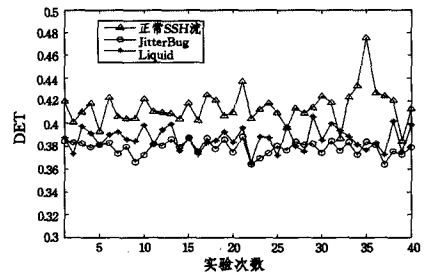

图 4　正常 SSH 数据流和各被动式隐蔽信道的 DET

表 2 为 100 组实验数据中各被动式隐蔽信道的检测率以及正常 SSH 信道的虚警率。从表 2 中我们可以看出，熵检测只能对 JitterBug 达到 100%的检测率，对 Liquid 只能达到 10%的检测率；而基于递归图的检测算法对 JitterBug 的检测率为 98%，对 Liquid 的检测率为 92%。这说明基于递归图的检测算法可有效应对两种被动式隐蔽信道的检测，而熵检测算法只能检测出 JitterBug，对 Liquid 却无能为力。

表 2　被动式隐蔽信道检测结果

检测阈值	SSH 虚警率	JitterBug 检测率	Liquid 检测率
EN<9.3	2%	100%	6%
CCE>1.82	4%	8%	10%
DET≤0.4	5%	98%	92%

4.3　检测窗口考察

为了考察检测窗口对检测率的影响，我们计算了不同检测窗口下的检测率，实验结果如图 5 所示。其中各窗口下的检测率是在 5%的误检率的前提下计算出来的。从图中我们可以发现 3 种主动式隐蔽信道在较小的检测窗口下也能达到很好的检测结果，当检测窗口为大于等于 800 时，基于递归图的检测算法对 IPCTC、TRCTC 和 MBCTC 3 种主动式隐蔽信道的检测率能够达到 100%。然而，被动式隐蔽信道却要依赖较大的检测窗口来实现较高的检测率，为了达到 90%以上的检测效果，被动式隐蔽信道需要的检测窗口大小为 2000，这点说明被动式隐蔽信道检测难度较大。

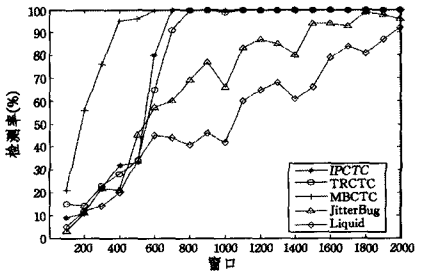

图 5　不同检测窗口下各隐蔽信道检测率

结束语　本文提出了一种基于递归图的检测方法，使用确定性(DET)作为递归图的量化分析指标，并把该指标用于检测隐蔽信道，得到了很好的检测效果。递归图反映了系统的相空间运动轨迹，本文从相空间轨迹的角度去考察正常 IPDS 和隐蔽信道的 IPDS 之间的区别，取得了很好的检测效果。事实上，无论主动式隐蔽信道还是被动式隐蔽信道，都对正常时间间隔进行大量的编码操作，必然改变了正常信道的相空间运动轨迹，使得隐蔽信道和正常信道之间在递归图特性上存在差别，因此使用递归图的量化分析指标可以有效地区分隐蔽信道和正常信道。

参 考 文 献

[1] Lampson B W. A note on the confinement problem[J]. Communications of the ACM, 1973, 16(10): 613-615

[2] Girling C G. Covert channels in LAN's[J]. IEEE Trans. on Software Engineering, 1987, SE-13(2): 292-296

[3] Cabuk S, Brodley C E, Shields C. IP covert timing channels: Design and detection[C]// Proc. of the 11th ACM Conf. on Computer and Communications Security. 2004: 178-187

[4] Cabuk S. Network Covert Channels: Design, Analysis, Detection, and Elimination[D]. Lafayette: Purdue University, 2006

[5] Gianvecchio S, et al. Model-Based Covert Timing Channels: Automated Modeling and Evasion[J]. Lecture Notes in Computer Science, 2008, 5230: 211-230

[6] Liu Guang-jie, et al. Covert Timing Channel with Distribution Match[C]// International Conference on Multimedia Informations Networking and security. 2009: 565-568

[7] Shah G, Molina A, Blaze M. Keyboards and Covert Channels [C]// Proceedings of the 15th conference on USENIX Security Symposium. 2006: 472-478

[8] Robert J Walls, Kush Kothari, Wright M. Liquid: A detection-resistant covert timing channel based on IPD shaping [J]. Computer Networks, 2011(55): 1217-1228

(下转第 149 页)

正常样本数/正常样本数。

半监督算法中不同的聚类质心的个数对聚类的效果影响是不同的，实验选取了不同的 K 值对上述 3 组数据分别进行了测试，取它们的平均值作为检测结果。图 1、图 2 给出了不同 K 值情况下云模型半监督聚类算法的检测结果。

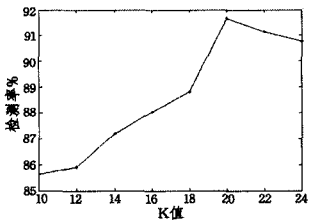

图 1　不同 K 值下的检测率

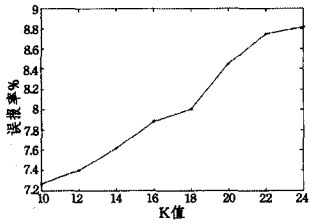

图 2　不同 K 值下的误报率

从实验结果可以看出，当 K 的值逐渐增大时，误报率也随之增大，但在 K 取 20 时，检测率获得最大，由此可知，K 取 20 时，基于云模型半监督聚类的算法可以获得较好的入侵检测效果，其检测率达到 91.67%，误报率为 8.45%。

在基于云模型的半监督聚类算法中，K 取不同值时的检测率与误报率与 K-means 算法的对比如图 3、图 4 所示。

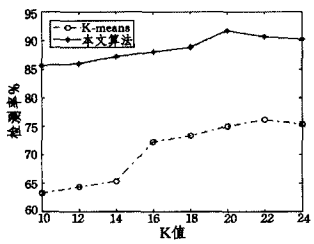

图 3　不同 K 值下检测率对比

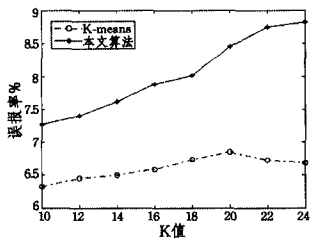

图 4　不同 K 值下误报率对比

由图 3、图 4 可以看出，在不同的 K 值情况下，本文算法在检测率方面明显高于 K-means 算法，误报率与 K-means 相比显得略微偏高，但是这种误报率在可接受的范围内。

基于云模型的半监督聚类算法检测结果与一般聚类算法以及普通云模型分类器[10]的比较结果如图 5 所示。

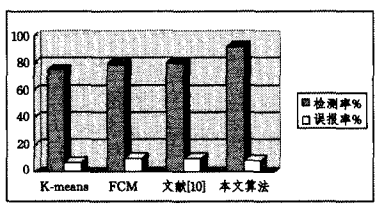

图 5　检测结果对比

图 5 给出了本文算法和其他几种算法的检测结果之间的比较，通过结果可以发现，本文算法的检测率明显高于其他 3 种算法，误报率比 K-means 略高，比其它两种算法的误报率低，证明了本文的算法具有优越的性能。

结束语　本文提出了一种基于云模型和半监督聚类的入侵检测算法，首先用改进的半监督的聚类对数据进行聚类处理，根据结果建立云模型，引入了云相对贴近度的概念，定义了高维空间样本在分类过程中的属性权重。在分类过程中对所建立的云模型更新和对属性实现动态加权不但能准确地反映实际数据信息而且指导了数据的分类，避免了对数据先验知识的依赖，在一定的程度上也丰富了云分类器相关的内容。实验证明了本文算法在入侵检测方面的可行性和有效性，但是本文算法的误报率仍然偏高，需要今后进一步的研究和改进。

参 考 文 献

[1] 李德毅，邸凯昌，李德仁，等. 用语言云模型发掘关联规则[J]. 软件学报，2000，11(2)：143-158

[2] 李德毅，史雪梅，孟海军. 隶属云和隶属云发生器[J]. 计算机研究和发展，1995，6(32)：15-20

[3] 李德毅，刘常昱. 论正态云模型的普适性[J]. 中国工程科学，2004，6(8)：28-34

[4] 吕辉军，王晔，李德毅，等. 逆向云在定性评价中的应用[J]. 计算机学报，2003，26(8)：1009-1014

[5] 付斌，李道国，王慕快. 云模型研究的回顾与展望[J]. 计算机应用研究，2011，28(2)：420-425

[6] 刘常昱，冯芒，李德毅，等. 基于云 X 信息的逆向云新算法[J]. 系统仿真学报，2004，16(11)：2417-2410

[7] Basu S, Baneriee A, Mooney R. Semi-supervised clustering by seeding[C]// Proceedings of the 19th International Conference on Machine Learning. San Francisco, CA: Morgan Kaufmann Publishers, 2002: 19-26

[8] Flanagan J A. Unsupervised clustering of symbol strings[C]// International Joint Conference on Neural Networks (IJCNN'03). Portland Oregon, USA: 2003, 3250-3255

[9] Li Yong-zhong, Li Zheng-jie. Anomaly Intrusion Detection Method Based on K-means Clustering Algorithm with Particle Swarm Optimization [C]// International Conference of Information Technology, Computer Engineering and Management Sciences(ICM 2011). 2006: 415-426

[10] 姜伟，高知新，李本喜. 基于多维云模型的入侵检测[J]. 计算机工程，2006，32(24)：155-156

[11] 李涵. 基于聚类的异常检测方法的研究与实现[J]. 北京信息科技大学学报，2010，25(3)：80-83

[12] KDD CUP 1999 Data set[OL]. http://kdd.ics.uci.edu/databases/ kddcup99

(上接第 117 页)

[9] Luo X, Chan E W W, Chang R K C. Cloak: A Ten-Fold Way for Reliable Covert Communications[C]// Proc. European Symp. Research in Computer Security. Sept. 2007

[10] Luo Xia-pu, et al. TCP covert timing channels: Design and detection[C]// IEEE International Conference on Dependable Systems and Networks With FTCS and DCC. 2008: 420-429

[11] Peng P, Ning P, Reeves D. On the Secrecy of Timing-Based Active Watermarking Trace-Back Techniques[C]// Proc. IEEE Symp. Security and Privacy. May 2006

[12] Berk V, Giani A, Cybenko G. Detection of covert channel encoding in network packet delays[R]. Technical Report, TR2005 536. Department of Computer Science, Dartmouth College, 2005: 1-11

[13] Gianvecchio S, Wang H N. Detecting covert timing channels: An entropy-based approach[C]// Proc. of the 14th ACM Conf. on Computer and Communications Security. 2007: 307-316

[14] Takens F. Detecting strange attractors in turbulence: Dynamical systems and turbulence[C]// Rand D A, Young L S, eds. Lecture Notes in Mathematics. 1981, 366

[15] Eckmann J P, Kamphorst S O, Ruelle D. Recurrence Plots of Dynamical Systems [J]. Europhysics Letter, 1987, 4: 973-977

[16] Groth A. Visualization of coupling in time series by order recurrence plots[J]. Phys. Rev. E, 2005, 72: 046220

[17] Waikato VIII[OL]. http://wand.net.nz/wits/waikato/8