第 35 卷 第 6 期
2013 年 6 月

系统工程与电子技术
Systems Engineering and Electronics

Vol. 35 No. 6
June 2013

文章编号:1001-506X(2013)06-1312-06

网址:www. sys-ele. com

基于校正熵的网络行为隐蔽信道的检测算法

钱玉文[1],宋华菊[2],赵邦信[1],张彤芳[1],郝劲松[1]

(1. 南京理工大学电子工程与光电技术学院,江苏 南京 210094;2. 南京晓庄学院,江苏 南京 210011)

摘 要:为解决传统检测方法检测隐蔽行为信道检测率较低的问题,提出了基于校正熵的隐蔽行为信道检测算法。所提算法利用向用户操作序列中嵌入隐蔽信息后,必然会引起其条件熵变化的原理进行检测。校正熵的引入,有效克服了利用条件熵进行检测会产生误报问题。基于校正熵对行为信道进行检测实验,检测结果表明,基于校正熵的隐蔽行为信道检测算法能够较好地在有噪声的环境中检测出几种常见的隐蔽时间信道,检测率约为 96%。

关键词:网络安全;隐写;隐蔽信道;熵

中图分类号:TP 393.08 **文献标志码**:A **DOI**:10.3969/j.issn.1001-506X.2013.06.31

Study on the detection algorithm of covert network behavior channel based on corrected entropy

QIAN Yu-wen[1], SONG Hua-ju[2], ZHAO Bang-xin[1], ZHANG Tong-fang[1], HAO Jin-song[1]

(1. *School of Electronic and Optical Engineering, Nanjing University of Science and Technology, Nanjing 210094, China*; 2. *Nanjing Xiaozhuang University, Nanjing 210011, China*)

Abstract: To solve the problem that traditional covert channel detection algorithms cannot detect covert behavior channel precisely, a detection approach based on corrected entropy is proposed. The idea of detection approach is that when embedding the information into network operations of the users, the condition entropy of some features of network operations would be changed. In order to solve the problem of false alarm of the detector based on comdition entropy, the detection approach based on corrected entopy is proposed. Several experiments are done to detect several covert behavior channels to get the performance of the detection algorithm based on corrected entropy. The detection results show that the algorithm can work well in detecting several covert behavior channels and the detection rate is about 96%.

Keywords: network security; steganography; covert channel; entropy

0 引 言

隐蔽信道作为一种严重危害网络安全的入侵,被定义为违背系统设计者原本意图,从系统的一个用户传送信息到另一个用户的机制[1]。网络隐蔽信道能够在监控的环境中窃取数据,造成严重的网络安全事故。网络隐蔽信道的检测问题已成为网络安全领域研究的热点。

隐蔽信道首先是在单机系统上开发的,最初对隐蔽信道的检测也是面向单机隐蔽信道的。安全系统中主要采用多级别安全机制,根据这种机制的特点,文献[2]提出了基于信息单向流动的检测算法。而文献[3]则根据隐蔽信道利用共享资源传输隐蔽信息的特点,提出了基于资源矩阵的检测算法,成为多安全级别系统中检测隐蔽信道的实用算法。但网络隐蔽信道根据网络数据流或数据包的特点来传输信息,不再依赖某个系统共享的资源,这往往使得传统的基于单机的隐蔽信道检测算法无法工作。

为了检测网络隐蔽信道,文献[4]提出了网络场景检测法,但实际网络中的场景过于复杂,这种检测算法的性能并不好。为了解决实际网络中隐蔽信道难以检测问题,研究者陆续提出了基于统计、学习的检测方法。例如,文献[5]采用基于 Kolmogorov-Smirnov (K-S) 的统计方法对隐蔽信道进行检测,文献[6-7]则使用 SVM 的隐蔽信道检测法,利用正常数据和异常数据的差异进行检测。然而,在实际的隐蔽通信中隐蔽信道产生的数据和正常通信产生的数据差异很

收稿日期:2012-06-15; 修回日期:2013-01-15; 网络优先出版日期:2013-06-07。

网络优先出版地址:http://www.cnki.net/kcms/detail/11.2422.TN.20130607.1835.012.html

基金项目:国家自然科学基金(61271230)资助课题

小，这类方法往往很难检测出隐蔽信道。正如文献[5]指出的，如果对特征量差别较小的数据进行检测，有可能会引起过学习，导致检测失败[5]。

在以上检测思想的启发下，一种新的网络隐蔽信道的检测思想逐渐形成，即如果正常的通信中加入隐蔽信息则必然给原来的通信数据带来某种改变。利用这种思想，文献[8]提出利用隐蔽时间信道的特点，通过波形的形状进行隐蔽信道检测的方法，这种方法很好地检测出时间信道。为了检测隐蔽时间信道中的未知规律，文献[9]提出了基于熵的时间信道检测算法。

网络隐蔽行为信道（简称为行为信道），利用用户使用网络时的行为来隐藏数据，这些行为往往表现为一系列的用户操作。由于熵只能衡量数据分布的混乱程度，它无法衡量操作序列的分布规律，因此，传统的基于熵的检测方法不适合用来检测网络行为信道。为了能够检测出网络行为信道，本文提出利用条件熵来衡量操作序列的规律性。为了克服条件熵在某些情况下可能出现的误判问题，引入基于校正熵的检测算法对网络行为信道进行检测。

1　网络隐蔽行为信道

按照调制隐蔽信息的原理不同，隐蔽信道一般分为隐蔽存储信道与隐蔽时间信道。隐蔽存储信道是一个进程直接或间接写一个存储位置，而另一个进程直接或间接地读该存储位置；而在隐蔽时间信道中，一个进程通过调制系统资源来传输信息给另一个进程。

早期的行为信道是利用网络包的某些时间行为来设计。这些网络包的时间行为，包括丢包行为、网络包到达次序错乱行为等。例如，文献[10-11]提出了基于包排序方法的行为信道，并指出在这种信道中，若发送 n 个网络包，则可以携带 $n!$ 个码字。随着隐蔽通信技术的发展，网络行为信道不再局限于网络包的时间行为，计算机网络中的多种行为，如登录服务器的顺序、访问服务器目录的顺序，浏览网站的行为等都可用作隐蔽行为信道的载体[12]。例如，文献[13]进行了长期的探索，先后提出利用 FTP 命令、HTTP 协议参数及多维 HTTP 参数等来设计隐蔽行为信道。因此，隐蔽行为信道既可以认为是一种存储信道，也可以认为是时间信道，它是新型的、具有较强隐蔽性的隐蔽信道[14]。随着时间的推移，服务器上的某些记录会消除，所以行为信道可能不会给取证者留下证据。由此可知，隐蔽行为信道具有暂态性、挥发性，这保证了它的隐蔽性。

行为信道与普通隐蔽信道的不同点表现在隐蔽信息被隐藏在某些正常的网络行为中，这些网络行为是一系列的用户操作。行为信道的设计者可以对不同的用户操作序列进行编码，隐蔽信道的接收者通过解码接收到的序列而获取隐蔽信息。若在某次网络通信中，对网络包的行为进行调制，则一定改变了原来用户操作的规律，新加入了一种规律。而且，只要有行为信道，便一定有一种新加入的行为规律，也就是说，增加新的行为规律是行为信道固有的特性，该特性可以用条件熵来衡量。

2　基于校正熵的检测算法

2.1　序列熵和条件熵

熵可以用来测量数据的复杂度和规律性，基于熵的检测方法是通过数据间的统计特性来判断是否存在异常情况。传统基于熵的检测方法只能检测数据的规律，对于数据之间的序列关系无法进行检测。然而，隐蔽行为信道往往使用多条命令、多个参数进行信息隐藏。为了检测数据间是否存在某些频繁出现的序列，需引入高阶随机的统计量来描述用户的网络行为规律。设用户的网络业务中共有 n 个操作，形成一个操作序列 $\boldsymbol{X}_n$。若用户在操作过程中共有 r 个操作可选择，则序列 $\boldsymbol{X}_n$ 中的每个操作都取自操作集合 $\boldsymbol{S}=\{x_1,x_2,\cdots,x_r\}$。这样，第 i 次操作为 $x_i(x_i\in S)$，则 x_i 的概率可表示为 $P(x_i)(i=1,\cdots,n)$，则用户的操作可看成一个随机过程 $\boldsymbol{X}_n$。将连续 m 个操作的概率写为向量 $\boldsymbol{X}_m$，可表示 m 维相空间上的一个点，在每一维上的取值分别为 $x_1,x_2,\cdots,x_m$[15]。这样，n 个用户操作便可分割成在 m 维相空间上点的集合，每个点表示一个行为的序列。此时，向量 $\boldsymbol{X}_m$ 可表示 m 维相空间上的一个点，该点可用来表示 m 个连续操作的一个行为模式。向量 $\boldsymbol{X}_m$ 的熵被定义为

$$\hat{H}(\boldsymbol{X}_1,\cdots,\boldsymbol{X}_m)=\sum_{x_1,\cdots,x_m}P(x_1)\times\cdots\times P(x_m)\log_2P(x_1)\times\cdots\times P(x_m) \tag{1}$$

由于 m 是一个正整数，而熵是在无限序列上计算出来的，因此，式(1)计算的其实是 $\boldsymbol{X}_m$ 熵的近似值，用 $\hat{H}$ 表示。将式(1)简写为

$$\hat{H}(\boldsymbol{X})=-\sum_m P_m\log_2P_m \tag{2}$$

式中，P_m 表示 $\boldsymbol{X}_m$ 的联合概率。向量 $\boldsymbol{X}_m$ 的熵表示在 m 维相空间中某一点 $\boldsymbol{X}_m$ 出现时的信息量。由此可知，若某些序列多次在熵空间中出现，则熵值较小。然而，熵仍然无法描述前一个事件发生对后一个事件的影响，而这种影响是用户操作行为的重要体现。因此，需要引入条件熵来描述用户的行为关系。实际上，由于 m 阶向量是在$(m-1)$阶向量的条件下，添加了一维向量而形成，因此条件熵可描述为

$$\hat{H}(\boldsymbol{X}_m\mid\boldsymbol{X}_1,\cdots,\boldsymbol{X}_{m-1})=-\sum_{m-1}P_{m-1}\sum_{m/m-1}P_{m/m-1}\log_2P_{m/m-1} \tag{3}$$

式中，P_{m-1} 表示 $\boldsymbol{X}_{m-1}$ 的联合概率；$P_{m/m-1}$ 表示条件概率，即第 m 个样本出现的条件是给定$(m-1)$个样本出现。式(3)中，第一项表示 $\boldsymbol{X}_m$ 在前$(m-1)$空间的熵，第二项表示在前$(m-1)$维相空间的基础上 $\boldsymbol{X}_m$ 的熵。由香农信息论可知，条件熵也可表示为

$$H(\boldsymbol{X}_m\mid\boldsymbol{X}_1,\cdots,\boldsymbol{X}_{m-1})=H(\boldsymbol{X}_1,\cdots,\boldsymbol{X}_m)-H(\boldsymbol{X}_1,\cdots,\boldsymbol{X}_{m-1}) \tag{4}$$

将式(4)写成序列的形式,即

$$H(m \mid m-1) = H(m) - H(m-1) \tag{5}$$

式中,$H(m)$为 $\boldsymbol{X}_m$ 在 m 维相空间上的熵;$H(m-1)$为 $\boldsymbol{X}_m$ 在$(m-1)$维相空间上的熵,表示条件熵。条件熵量化了指定一个新的状态在一位增量空间的信息。当长度为 m 的行为模式可完全由前$(m-1)$维的行为模式决定时,则条件熵为 0。反之,若样本向量 $\boldsymbol{X}_m$ 的各维分量相互独立,则 $\boldsymbol{X}_m$ 中第 m 个变量概率 P_m 可由链接概率计算,即 $P_m = P_1 \times P_{m-1}$。因此,当向量 $\boldsymbol{X}_m$ 中各个分量相互独立时,$\boldsymbol{X}_m$ 的熵为

$$H(m) = -\sum_m P_1 P_{m-1} \log_2 P_1 + \sum_m P_1 P_{m-1} \log_2 P_{m-1} \tag{6}$$

将一维相空间展开写成求和的形式,即

$$\hat{H}(m) = -\sum_1 \sum_{m-1} P_1 P_{m-1} \log_2 P_1 + \sum_1 \sum_{m-1} P_1 P_{m-1} \log_2 P_{m-1} \tag{7}$$

考虑到 $\sum_1 P_1 = \sum_{m-1} P_{m-1} = 1$,则

$$\hat{H}(m) = -\sum_m P_1 \log_2 P_1 - \sum_{m-1} P_{m-1} \log_2 P_{m-1} = H(1) + H(m-1) \tag{8}$$

根据式(5)有

$$\hat{H}(m \mid m-1) = H(m) - H(m-1) = H(1) \tag{9}$$

由此可知,$H(m)$可以通过计算 $H(m-1)$和 $H(1)$获得。结合式(5)和式(8),若样本独立,这表示样本间的各个命令的出现是随机的(白噪声的情况),条件熵便是一个常数。由此可知,若没有固定的序列出现则熵较大。反之,如果是一个严格的周期过程,则通过$(m-1)$维相空间的样本便可以预测第 m 维样本的值,条件熵为 0。因此,可以利用条件熵来预测用户的某些操作序列是否多次出现,从而来预测是否存在潜在的隐蔽信道。这是因为用户若使用序列来嵌入信息,由于需要对序列进行编码,则某些序列可能会频繁出现。因此,可以使用条件熵来预测序列出现的频度。

式(9)描述的是一种特殊的情况,即 m 维数据中各维都独立时的情况。在各维不独立的情况下,条件熵只能使用式(10)来估算。

$$\hat{H}(m \mid m-1) = \hat{H}(m) - \hat{H}(m-1) \tag{10}$$

式中,$\hat{H}(m)$,$\hat{H}(m-1)$分别为 m 维及$(m-1)$维相空间上向量 $\boldsymbol{X}_m$ 的熵的估计值。假设在$(m-1)$维相空间中某个点只出现一次,将该点代表的向量再增加一维形成了 m 维相空间中的一个点,则可以确定在 m 维相空间中该点只会出现一次。因此,m 维相空间中出现单点的情况可完全由$(m-1)$维来确定,概率为 1,即$(m-1)$维相空间中出现了某个单点,则将延伸到 m 维后,对应的点仍然为单点是一个确定的事件。因此,这种点对条件熵的估计值 $\hat{H}(m|m-1)$没有任何贡献。由于维数 m 越大,单点的个数会越多,则 $\hat{H}(m|m-1)$的值会越来越小。这就意味着维数越大,序列的随机性越大,而条件 $\hat{H}(m|m-1)$越小。当一个数据序列是完全随机的,则条件熵为 0。然而,条件熵为 0 的情况表示数据间的规律性最强,最有可能存在隐蔽信道,这种决策显然是错误的。

另外,如果使用条件熵作为检测的依据,选择行为模式长度是否合适,会对检测算法的性能影响极大。如果行为模式长度选择的不合适,会出现预测错误的现象。例如,一个行为序列的长度为 5,则若选择序列的长度小于 5 时,行为就无法显现出来。但选择的行为模式长度过大时,在检测过程中就会出现单点的情况急剧增加,而使得条件熵减小,会产生误导检测器的情况。因此,若使用条件熵来检测隐蔽信道,行为模式长度的选择会成为用户很大的负担。为了克服使用条件熵进行决策时产生误报的问题,以及行为模式长度很难选择的问题,本文引入校正熵的方法对行为信道进行检测。

2.2 校正熵

为了解决检测过程中单点对条件熵的影响,以及行为模式长度选择的问题,首先,需要了解在 m 维空间中熵的计算。设在 m 维相空间中,由于某个点在某个位置单独出现的事件是一个随机过程,使用单点熵来衡量 m 维空间出现单独点的平均互信息。相似地,使用多点熵来衡量某个序列在 m 维空间中出现 2 次以上点的平均互信息,则在 m 维相空间中序列 X_m 的熵为出现单点的熵与出现多点的熵之和,表示为

$$H(m) = H_1(m) + H_2(m) \tag{11}$$

式中,$H_1(m)$为单点的熵;$H_2(m)$为多点的熵。由信息论的知识可知,单点的熵可表示为

$$H_1(m) = P(m)\log_2(n-m+1) \tag{12}$$

式中,$P(m)$是单个点出现的概率;n 是样本总数。单点的熵随着 m 的增大而增大,而多点的熵随着 m 的增大而减小。由于条件熵在检测时,单点的条件熵为 0,但并非意味着存在一种行为模式。因此,需要将单点熵值加入到衡量是否有行为模式出现的检验标准中,以此来弥补条件熵的减小值。基于以上考虑,本文提出使用式(13)来改进条件,作为检验是否存在隐蔽行为信道的判据。

$$KH(m \mid m-1) = \hat{H}(m \mid m-1) + H_c(m) \tag{13}$$

式中,$H_c(m)$根据式(12)来定义,可表示为

$$H_c(m) = P(m) \times \hat{H}(1) \tag{14}$$

式中,$\hat{H}(1)$是 $m=1$ 时的熵值,$\hat{H}(1)$表示当白噪声的序列出现时,校正熵 KH 的取值;$P(m)$是单个点出现的概率。利用这种方法校正后,若没有可靠的规律被计算出来,即数据随机性较大,校正熵的值接近于 1 维情况下的熵值,不会降至很小。在式(13)中,由于校正的条件熵 KH 有两项组成,一旦第一项减少,第二项便会增加。如果隐蔽行为信道中,有一些序列经常出现,则校正熵会取一个最小值,这个最小值便是 KH 的最佳预测点。若最小值与正常值的差别较小,则认为没有隐蔽行为信道存在,否则便认为存在隐蔽信道。利用校正熵可以有效克服利用条件熵在数据随机分

布时错误决策的问题，下面通过设计基于校正熵的检测算法来解决长度选择的问题。

2.3　基于校正熵的检测算法

在检测过程中，将 m 维的量化空间定义为一个 m 维的超立方体，m 维空间的每个点都表达了一个长度为 m 的模式。若相同的点出现多次，则可能存在隐蔽行为信道。事实上，即使是相同的序列出现在同一个点的可能性也是非常小的，它们往往是一个相近的值，即由相同的序列产生的 m 维向量上的点距离较近。因此，为了预测序列出现的频率，需将各个序列出现的概率进行量化处理。引入量化因子 ξ，将序列 $\boldsymbol{X}_i$ 分散到 ξ 个量化级别中，则量化的单位元为 $\varepsilon=(\boldsymbol{X}_{\max}-\boldsymbol{X}_{\min})/\xi$。量化处理后，超立方体内的所有点之间的距离都小于 ε。设在超立方体内，两个点的误差小于阈值 e，则表示它们的模式相同。反之，若某个点只出现一次，则它代表的模式只出现一次。

为了计算校正熵，需估计有限序列的条件熵与熵。若操作命令的个数最多为 q 个，选择一个操作序列最大可能的长度（如 10，表示一个操作行为由 10 个命令组成）。将此最大序列长度记为 l，则可利用 l 的 q 叉树来计算校正熵。而在该树中计算得到的最小校正熵将会用来作为判断是否有隐蔽行为信道的依据，计算最小校正熵的算法如算法 1 所示。

算法 1　建立行为模式树及计算最小校正熵的算法

步骤 1　选择根操作，将根操作进队列 Q。

步骤 2　判断队列 Q 是否为空，若队列为空则退出，否则转到步骤 3。

步骤 3　将队列 Q 中的队头出队列，得到当前操作。从所有样本中挑选出包含当前操作的序列，并选择当前操作的后一个操作。将序列中的所有操作入队列 Q。并将所有这些操作添加到模式树中，在操作时由于共有 q 个操作，则每个节点最多有 q 个子树，并且加入时操作按子树进行编号。

步骤 4　计算确定 q 个操作的概率，即计算每个子树出现的概率。

步骤 5　判断是否还有样本数据，如果有，转到步骤 2，否则执行步骤 6。

步骤 6　对建成的模式树进行广度优先遍历，广度优先遍历可以计算各个节点的校验熵，相加后形成各层的条件 KH，记为：$KH[1],KH[2],\cdots,KH[m]$，并从中取得最小值，即

$$KH_{\min}=\min\{KH[1],KH[2],\cdots,KH[m]\} \quad (15)$$

通过算法 1 的方法建立的行为模式树中，从根结点到叶子结点的一条路径便是一个行为模式，路径的长度便是行为模式的长度，为 l。通过行为模式树可以方便的计算各个行为模式的操作序列出现的次数。在建立行为模式树时，增加一个新的长为 l 的行为模式时，向叶子节点前进，更新中间节点的个数，同时建立新的节点。当遍历到达树的最后一层时，已经将新的行为模式以及它所有的子行为模式都统计一遍。这样，只要将 l 设为一个较大的值，便能通过 $KH_{\min}$ 确定适当的行为模式长度。一般而言，行为模式的长度为 10，若超过 10，隐蔽信道所传输的信息将会非常小，这种信道的意义较小，随着模式长度的增加检测时间也急剧增大。

3　检测实验及结果分析

3.1　基于 FTP 命令的行为信道的检测实验

为了检测不同类型的隐蔽行为信道，需要对正常的隐蔽信道数据以及带有行为信道的隐蔽信道的通信数据进行训练，训练的结果是发现其中校正熵的差异。选择基于 FTP 协议的通信数据作为实验样本，即利用 FTP 的协议行为来调制隐蔽信息。首先，总结并提取常用的几种 FTP 的操作序列作为调制对象，每个操作序列由若干条命令组成；其次，对这些命令序列进行编码，将出现概率大的命令映射成权值大的数据，否则映射成权值小的数据；第三，对这些数据进行 haffman 编码，其结果放置在 haffman 树中。选择常用的几种 FTP 的行为：下载文件行为、上传文件行为、浏览目录等作为实验的对象，进行嵌入数据。每种行为的网络包个数为 1.2 万个。计算以上各种行为在不同的行为模式长度下校正熵的值，如图 1 所示。

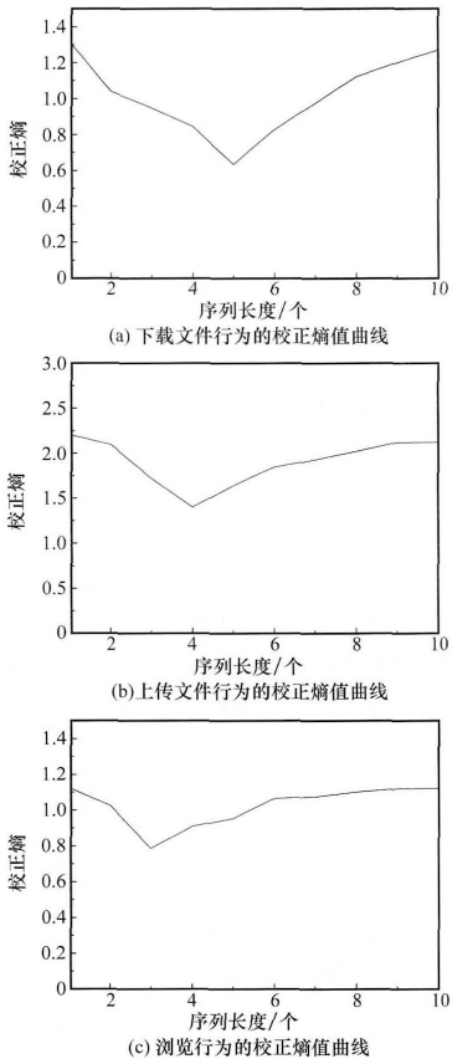

图 1　不同 FTP 行为的隐蔽信道在有噪声环境下的校正熵值曲线

图 1 中，选择的数据中含有 1.2 万条操作。这些数据并非完全是带有 FTP 行为的数据包。这样其校正熵不会为 0。但是，由于噪声的比例较小，因此，在选择较为合适的行为模式长度后，校正熵的值会非常小。图 1 中，当含有隐蔽行为信道的包占所有包的 40% 时，校正熵的取值最小约为 0.7。而随着序列长度的增大，校正熵的值越来越接近于随机情况下 1 维数据的熵。

为了能够检测隐蔽行为信道，还需要检测 FTP 命令为正常所使用时，校正熵的取值情况，选择在一台服务器上某一时段内 1.2 万条 FTP 命令，并且去除其中一些固定的用户操作序列，然后计算其校正熵，如图 2 所示。

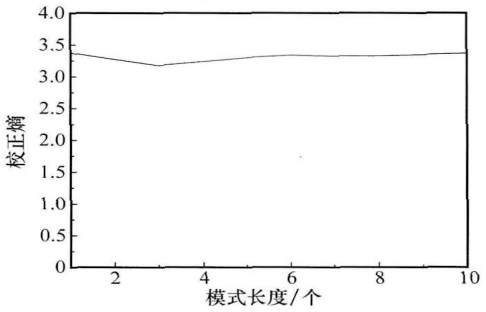

图 2 正常 FTP 行为的隐蔽信道在有噪声环境下的校正熵值曲线

由图 2 可知，由于各个命令出现的情况较为相似，操作的序列表现出的规则性较小，因此出现的单点情况较多，此时校正熵的变化范围不大。基于这种考虑，选择校正熵的最小阈值为 0.7，即检测是每次选择 3 000 个网络包作为检测窗口，如果发现利用式(14)计算出的阈值小于 1.5，即认为存在隐蔽行为信道，否则认为没有隐蔽行为信道。

选择 20 万个 FTP 网络包，这些网络包是实际 FTP 服务器上的记录。选择不同的含有隐蔽信道行为命令的网络包进行检测。为了衡量含有隐蔽信道的数据在正常数据中的比率，定义行为噪声比为

$$\text{行为噪声比} = \frac{\text{不含有隐蔽信道命令的数据包}}{\text{网络包的总数}} \quad (16)$$

向正常的网络数据中注入不同比例的行为噪声，使用算法 1 的方法对这些数据进行检测，结果如图 3 所示。

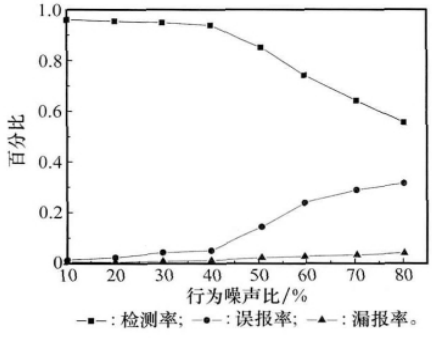

图 3 FTP 行为信道在具有噪声的实验环境下的检测率

由图 3 可知，行为噪声比较小时，该方法可有效地发现隐蔽信道。行为噪声比加大时，检测精度略微有所改变，但行为噪声比大于 40% 时，检测器的检测效率会急剧下降。行为噪声比大于 40% 时，以 1.5 为阈值很难检测到隐蔽信道。因此，阈值的选择对检测效果影响较大。行为噪声比越大，则数据的随机性便越大，校正熵也越大，此时需要选择的阈值便较大，否则误报率与检测精度便会急剧下降。然而，行为噪声比与校正熵之间取值的关系较为复杂，它受数据随机程度以及样本数量等因素影响，无法非常精确地对应关系式。目前，一般只能通过数据训练、实验等方法获得在某种级别的行为噪声比下校正熵的取值。在检测过程中，当选择检测窗口为 3 000 个网络包，模式长度为 10 时，检测时间约为 2.8 s，该时间是可以被网络管理员所接受的。

3.2 检测结果比较

由以上实验可知，基于校正熵的行为信道检测方法，可以在有噪声的情况下，较好地检测出基于 FTP 命令设计的行为信道。下面进行第 3 个检测实验，在该实验中利用基于校正熵的检测算法与两种较好的隐蔽信道检测算法(即 ε 打分法[8]，规则检测法[9])分别对 3 种隐蔽信道进行检测，比较其检测结果。

为了让仿真数据具有一定的典型性，选择某学校校园网内的 FTP 服务器上 1 个月的数据作为正常的数据，并使用当前较为常见的行为信道产生异常通信的数据，这些行为信道包括：①利用 FTP 使用的命令集合构成的行为信道；②利用 HTTP 命令(post，get 等)构成的隐蔽信道[16]；③利用 TCP 参数构成的行为信道[12]；④针对包排序的行为信道[11]。为了测试检测过程的实时性，设定检测窗口值为 3 000个数据包。在实验中，发送主机分别使用不同的隐蔽信道的发送数据。发送时的场景是：①3 种隐蔽信道分别由 1 个主机向接收端发送隐蔽数据；②含有 3 种隐蔽信道数据的网络包各有 2 个；③在接收端安装具有几种检测算法的检测器；④从公共数据集中选择网络包作为噪声，加入到发送队列中进行发送，将这 3 种信道的行为噪声比控制在小于 20% 的条件下。从接收端记录检测结果，如表 1 所示。

表 1 不同隐蔽信道检测表 %

方法	HTTP 命令行为信道		TCP 参数行为信道		包排序信道	
	检测率	误报率	检测率	误报率	检测率	误报率
ε 打分法[7]	85	17	86	20	88	17
规则检测法[8]	80	11	76	14	84	12
基于校正熵	96	7	95	11	98	5

在检测过程中，使用校正熵对表 1 所列的 3 种隐蔽信道分别选择阈值为 1.5，2.2 和 1.4。由表 1 可知，传统的检测算法对于隐蔽行为信道几乎无法检测，因为传统的检测算法没有关注到各个网络包之间的关系，即用户操作的序列关系形成的网络包之间的关系[17]。而基于密度聚类的方法对该种隐蔽信道的检测率平均达到 96.1%，误报约率

为 9%。因此，这种算法能够较好地检测出网络行为隐蔽信道。

4 结　论

本文分析了基于校正熵的隐蔽行为信道的检测原理。通过校正熵方法对已被报道的几种隐蔽行为信道进行检测。由于用户只要将信息调制到特定的网络行为中，必然改变其数据的统计规律，从而引起条件熵的变化，因此这种方法可检测多种隐蔽行为信道，具有一定程度上的盲检测的功能，且检测率与实时性较好。然而，实际的检测过程中，校正熵的值随着训练数据的数量、质量的变化而变化，这也给用户在检测时选择合适的阈值带来一定的困难，如何使得检测器能够自适应地选择阈值也是本课题下一步研究的方向。

参考文献：

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

[2] Denning D E. A lattice model of secure information flow[J]. *Communications of the Association for Computing Machinery (ACM)*, 1976, 19(5):236-243.

[3] Kemmerer R A. Shared resource matrix methodology: an approach to identifying storage and timing channels[J]. *ACM Trans. on Computer Systems*, 1981, 1(3):256-277.

[4] Helouet L, Jard C, Zeitoun M. Covert channels detection in protocols using scenarios[C]// *Proc. of the Workshop on Security Protocols Verification*, 2003: 21-25.

[5] Murdoch S J, Lewis S. Embedding covert channels into TCP/IP[C]// *Proc. of the 7th Information Hiding Workshop*, 2005: 6-8.

[6] Sohn T, Moon J, Lee S, et al. Covert channel detection in the ICMP payload using support vector machine[C]// *Proc. of the International Symposium on Computer and Information Sciences*, 2003: 828-835.

[7] Qian Y W, Song H J, Song C. Network covert channel detection with cluster based on hierarchy and density[J]. *Procedia Engineer*, 2012,29:4175-4180.

[8] Sendar C, Carla B, Clay S. IP covert channel detection[J]. *ACM Trans. on Information and System Security*, 2009, 12(4): 221-227.

[9] Gianvecchio S, Wang H. Detecting covert timing channels: an entropy based approach[C]// *Proc. of the ACM Conference on Computer and Communications Security*, 2007: 307-316.

[10] Ahsan K, Kundur D. Practical data hiding in TCP/IP[C]// *Proc. of the ACM Workshop on Multimedia Security*, 2002: 72-77.

[11] Ji L P, Fan Y, Ma C. Covert channel for local area network[C]// *Proc. of the International Conference on Networking and Information Security*, 2010: 316-319

[12] Wang Y J, Wu J Z, Zeng H T, et al. Covert channel research[J]. *Journal of Software*, 2010, 21(9): 2263-2288. (王永吉，吴敬征，曾海涛，等. 隐蔽信道研究[J]. 软件学报，2010, 21(9): 2263-2288.)

[13] Zou X G. Covert channels based on command sequence of FTP protocol[J]. *Journal of Harbin Institute of Technology*, 2007, 39(3):424-426.(邹昕光. 基于 FTP 协议的命令序列隐蔽信道[J]. 哈尔滨工业大学学报，2007, 39(3):424-426.)

[14] Jason J, Ridha K, Zhang Q L. On the necessary conditions for covert channel existence: a state-of-the-art survey[C]// *Proc. of the International Conference on Ambient Systems*, 2012: 458-465.

[15] Packard N H, Crutchfield J P, Farmer J D, et al. Geometry from a time series[J]. *Physical Review Letter*, 1980, 45(9): 712-716.

[16] Qian Y W, Zhao B X, Kong J S, et al. Robust covert timing channel based on Web[J]. *Journal of Computer Research and Development*, 2011, 48(11):423-431.(钱玉文，赵邦信，孔建寿，等. 一种基于 Web 的可靠网络隐蔽时间信道的研究[J]. 计算机研究与发展，2011,48(11):423-431.)

[17] Zhu Y, Yu M Y, Hu H X, et al. Efficient construction of provably secure steganography under ordinary covert channels[J]. *Science China Information Sciences*, 2012, 55(7): 1639-1649.

作者简介：

钱玉文(1975-)，男，讲师，博士，主要研究方向为网络安全、信息隐藏等。

E-mail: admon1999@163.com

宋华菊(1978-)，女，讲师，硕士，主要研究方向为信息化学、信息处理、计算化学。

E-mail: songhuaju2011@163.com

赵邦信(1966-)，男，副教授，主要研究方向为信息论与编码、计算机测控技术。

E-mail: zhaobx@mail.njust.edu.cn

张彤芳(1989-)，女，硕士研究生，主要研究方向为网络通信技术。

E-mail: tongfangzhang@gmail.com

郝劲松(1987-)，男，硕士研究生，主要研究方向为图像处理、数据挖掘。

E-mail: haojinsong@163.com