基于小波变换的木马心跳行为检测方法

白 虹 庞建民 戴 超 岳 峰
(数学工程与先进计算国家重点实验室 郑州 450001)

摘 要 通常的木马心跳行为检测方法利用的是聚类的思想,很难避免木马自身传输数据包的干扰,导致误报。为此,提出基于小波变换的木马心跳行为检测方法。该方法首先将TCP数据包流表示成包长度信号,然后用基于Mallat的强制阈值除噪算法对信号进行处理,最后通过基于包速率的行为详细信息判定算法得出检测结果。实验表明,该检测方法能有效地检测出心跳行为并具有较强的抗干扰性。

关键词 木马心跳行为,包长度信号,Mallat定理,小波变换

中图法分类号 TP393.08 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2016.4.030

Trojans Keep-alive Behavior Detection Approach Based on Wavelet Transform

BAI Hong PANG Jian-min DAI Chao YUE Feng
(State Key Laboratory of Mathematical Engineering and Advanced Computing, Zhengzhou 450001, China)

Abstract Trojans keep-alive behavior detection algorithms generally are based on the method of clustering, which can hardly avoid the interference of other packets in the network, leading to false positive results. Therefore, this paper proposed a Trojans keep-alive behavior detection approach based on wavelet transform. In this approach, firstly, TCP packets stream is described by packet length signal, then the signal is processed by compelling threshold denoising method based on Mallat theory, and finally detection results can be acquired through detail information decision algorithm based on packet rate. Experiments show that this approach can detect Trojan keep-alive behavior effectively and has better anti-interference.

Keywords Trojans keep-alive behavior, Packet length signal, Mallat theory, Wavelet transform

1 引言

近年来,恶意代码的数量增长惊人,对网络安全构成越来越大的威胁。其中木马程序占有重要比例,并且引发网民隐私泄露、财产安全以及机构安全的问题。研究人员通常通过检测并查杀木马来解决这一问题。目前木马的检测方法主要是在主机上检测其特征码,但是对未知的木马或者已知木马的变种则会失效。Scarfone K 和 Mell P[1]提出恶意代码网络行为检测方法,该方法是指在网络流量中识别恶意代码生成的特殊流量,比如木马的心跳行为、控制行为等。虽然木马程序种类繁多并且有大量变种,但木马程序的网络行为却是有限的,所以基于网络行为的木马检测方法能有效检测出未知木马。正因如此,木马的网络行为研究开始受到研究者的关注。丁卫球[2]提出一种基于进化矩阵的木马行为特征检测方法,其对特征流量和非特征流量加以区分,从而识别出各种木马。夏爱民、张宏志和杨伟等人[3]研究了Windows系统,基于木马综合行为特征分析的木马检测技术,提出木马检测算法,该算法能有效识别木马。马立军[4]提出基于行为检测的窃密型木马检测方案,通过对常见窃密型木马通信机制进行建模分析,构建窃密型木马的几种通信模式。

然而心跳行为是最重要的木马网络行为之一。心跳行为不仅是木马的必须行为,还能表明木马存活与否。因而,木马心跳行为检测是不可或缺的。目前木马心跳行为的主要检测方法都是应用聚类的思想,对数据包先分簇再处理来识别心跳行为的特征。Tao He 和 Hao Zhong[5]提出一种周期性序列的检测技术,这种技术结合了模糊原理,将网络数据包中抽取的信息保存在数据库中,并把网络数据包聚类成序列,然后通过计算聚类序列的相似度和阈值检测出心跳包。该方法能在网络流中较准确地识别心跳包,但并不高效。易军凯等[6]提出一种基于数据流分簇处理的检测心跳包序列的方法。这种方法对数据流进行时序分簇处理,按周期性及其他特征扩充或筛除簇集合来检测心跳行为。该方法能够进行实时检测,但抗干扰能力较弱。Yiguo Pu[7]等人将心跳行为分为3类,分别是TCP协议的keep-alive机制、基于TCP短会话的心跳和基于TCP长会话的心跳,并且设计了相应的检测方法。检测时使用3种方法一一排除,可以减少漏报。但是,其中基于TCP短会话心跳的检测方法是按数据包大小聚类所有的包,然后判断每一类中是否有周期性。这一方法忽略了每次心跳行为的相应数据包的IP需相同这一特征,很可能产生误报。另外,网络流量中常常混有相同IP的非心跳包,例

到稿日期:2015-03-30 返修日期:2015-06-13 本文受国家自然科学基金:基于仿生学原理的恶意代码判定与防护模型研究(61472447)资助。
白 虹(1990-),女,硕士生,主要研究方向为信息安全,E-mail:baihong00@163.com;庞建民(1964-),男,教授,主要研究方向为高性能计算、信息安全;戴 超(1982-),男,博士生,主要研究方向为信息安全;岳 峰(1985-),男,讲师,主要研究方向为系统虚拟化。

如木马操作屏幕截图这一功能时，在与心跳包相同的 IP 端之间会有大量的其他数据包进行传输截屏。由于短时涌现大量数据包，若用上述方法检测心跳，则会产生很多数据包的聚类，从而影响心跳包的检测。

此外，木马心跳行为检测还包括采取信号处理的方法。孟磊、刘胜利等人[8]通过使用时频分析法分析大量样本，发现正常网络通信数据包的时间间隔通过离散傅里叶变换到频域后，表现出平稳特性，而木马表现出非平稳信号的特性，中高频系数几乎为 0。因此，他们通过判断心跳间隙的中频和高频系数的值域来实现检测。但是该方法没有给出心跳行为的形式化描述，并且离散傅里叶变换的方法缺乏局部性，只能计算单个频谱，不能反映出信号在某时间段的局部特性。因为网络中的流量信号是随时间变化的，使用具有局部性的小波分析更能反映出信号的特征。基于以上问题，本文提出一种基于小波变换的木马心跳行为检测方法。该方法首先定义了心跳行为的信号表示，并把网络数据包流表示成信号，然后对此信号进行基于小波变换的强制阈值去噪算法处理得到近似心跳信号，最后通过详细信息判定近似心跳信号是否为心跳行为。本文首先介绍该方法的原理，然后陈述具体算法的设计，最后给出实验过程及结果。

2 木马心跳行为检测算法原理

网络流量常常由各种数据包组成，其中包含反映木马心跳行为的数据包，也包含其他数据包。因此，需从混有各种数据包的网络流量中检测心跳行为数据包。这与医学中的生物电信号检测相似。当研究者试图采集人的眼电信号时，其中一定会混有肌电信号和脑电信号等其他生物电信号[9]。然而，医学工作人员通过对所采集的信号先进行多级去噪再进行分析的方法可以有效检测某种生物电信号。因此，本文将医学生物电信号处理的思想应用于木马心跳行为检测，提出木马心跳行为检测算法。本节首先定义包长度网络信号，然后介绍木马心跳行为的特征，最后详细描述心跳行为检测算法的原理。

2.1 包长度网络信号的刻画

由于生物电信号处理的对象是信号，因此使用信号处理的方法检测木马心跳行为需要将网络流量或数据包刻画成信号，以便进行分析。信号多用在通信领域，定义 1 是奥本海姆给出的最初的信号定义。

定义 1(信号)　信号可以描述范围极为广泛的一类物理现象。信号所包含的信息总是寄寓在某种变化形式的波形之中。在数学上，信号可以表示为一个或多个变量的函数[10]。

在网络方面，大部分文献都是用“流量信号”刻画网络中的流量，具体描述见定义 2。例如，杨岳湘建立流量信号[11]，对网络流量进行研究；杨继鹏通过对流量信号进行小波分解[12]来研究异常检测。

定义 2(流量信号)　自变量为时间，因变量是流(包、字节)数的函数。在坐标轴上表示即为横轴是时间，纵轴是流(包、字节)数。

然而，流量信号适合应用在网络流量数量会骤然增加、有流量突变的行为的检测，如 DDoS 攻击、流量拥塞等，对木马心跳行为检测则没有效果。因此，本文提出适用于木马心跳检测的包长度网络信号，详细描述见定义 3。

定义 3(包长度网络信号)　自变量为时间、因变量是数据包长度的函数表示包长度网络信号，记作 $u(n)$。映射到坐标轴上即横轴为时间，纵轴为数据包长度，如图 1 所示，使用包长度网络信号描绘了网络中的一段流量。

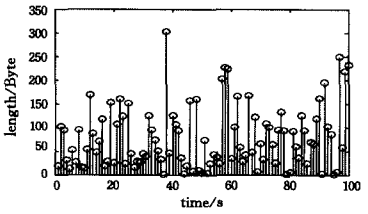

图 1　数据包长度信号表示的一段网络流量

2.2 木马心跳行为特征

木马建立连接后，木马通信两端为继续维持连接状态，服务端与客户端会周期性地相互发送数据包，这种表征木马存活的数据包称为木马的心跳包。这种周期性地发送心跳包的行为通常称作木马的心跳行为。其可由 TCP 协议自带的 keep-alive 机制实现，也可以由应用层实现，但是都具有如下特征。

(1)交互性与一致性。心跳包通常是由 2 个以上(包括 2 个)的数据包组成一组来完成一次心跳过程，类似 TCP 的握手连接。同一款木马每次的心跳包的组合都是一致的。

(2)数据包负载小。通常的木马设计心跳包只作为应答包，所以其负载较小。可能有个别的木马会增加心跳包的负载，在这里不做深入研究。

(3)周期性。通常传统的木马心跳周期都是固定的，并且周期明显比一般数据包的时间间隔长，有区分度。现在有一些木马是变频的，暂时不作考虑。

2.3 木马心跳行为检测算法原理

在检测木马的网络环境中，不会只有心跳数据包，还会有大量的其他数据包混合在一起，使得心跳信号难以觉察。若把非心跳数据包当作噪声，则网络信号可以认为是心跳信号与噪声信号的混合，可用式(1)表示。

$$x(n)=s(n)+u(n) \tag{1}$$

其中，$x(n)$是有心跳的网络数据包信号，$s(n)$是噪声信号，$u(n)$是心跳信号。噪声信号 $s(n)$是离散随机序列，是非平稳信号。心跳信号 $u(n)$是离散周期信号，是平稳信号，可以用式(2)表示。

$$u(n)=a\delta(n+kT)+b\delta(n+kT-1) \tag{2}$$

其中，a、b 为常数，$\delta(n)$为单位抽样信号。由于木马心跳的周期性较长，相比噪声信号，可以认为信号 $u(n)$是混合信号 $x(n)$中的低频部分，噪声信号是高频部分。图 2 分上下两个子图，位于上部的信号图描述了心跳行为，下部的图是心跳行为信号的频谱图。图 3 描绘了噪声信号和噪声信号的频谱图。图 4 是图 3 噪声信号的局部放大图，与图 2 比较，可以得出噪声信号的频谱比心跳信号的频谱大很多。因此可以通过一些变换或算法滤除混合信号的高频部分，从而剩下心跳信号，以方便检测心跳行为。

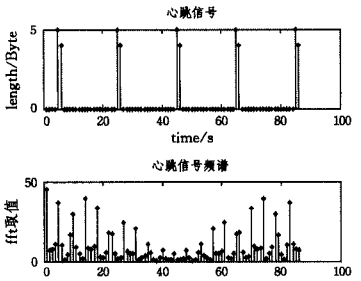

图 2　心跳信号及其频谱

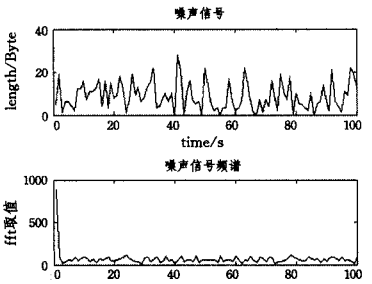

图 3　噪声信号及其频谱

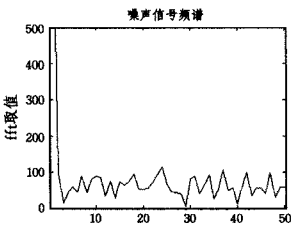

图 4　噪声信号频谱局部放大图

3　木马心跳行为检测算法

根据木马心跳行为检测原理，设计了基于 Mallat 的强制阈值除噪算法，对原混合信号进行处理得到近似心跳信号，然后再通过详细心跳行为识别算法得出判断结果。下面分别介绍这两个算法。

3.1　基于 Mallat 的强制阈值除噪算法

根据第 2 节中所介绍的原理可知，对网络信号进行除噪就是要消除网络信号中的高频部分。小波分析可以将信号分解成高频部分和低频部分，并且具有多分辨率分析的特点，可以对网络信号进行多层分解，有效地将低频信号与高频信号进行分离。Mallat 算法是小波分析中的经典算法，为分离出低频信号得到近似心跳信号，设计了基于 Mallat 的强制阈值除噪算法。

3.1.1　Mallat 定理

1988 年 S. Mallat 给出了正交小波的构造方法以及正交小波的快速算法，即 Mallat 算法。Mallat 算法在小波分析中的地位相当于快速傅里叶变换算法在经典傅里叶分析中的地位。Mallat 算法的分解过程如图 5 所示。

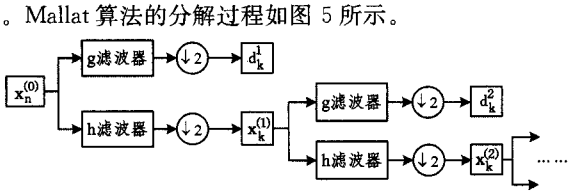

图 5　Mallat 分解过程示意图

其中，$x_k^{(j)}$ 为离散平滑逼近，d_k^j 是离散细节信号（也称细节系数），$\downarrow 2$ 表示二抽样，即从 $x_k^{(j)}$ 到 $x_k^{(j+1)}$ 和 d_k^{j+1} 样点数减少一半。g 是高通滤波器，通过 g 得到信号的高频部分；h 是低通滤波器，通过 h 得到信号的低频部分。具体计算如式(3)—式(6)所示。

$$x_k^{(1)}=\sum_n h(n-2k)x_n^{(0)} \tag{3}$$

$$d_k^1=\sum_n g(n-2k)x_n^{(0)} \tag{4}$$

往下逐级引申，有以下公式：

$$x_k^{(j)}=\sum_n h(n-2k)x_n^{(j-1)} \tag{5}$$

$$d_k^j=\sum_n g(n-2k)x_n^{(j-1)} \tag{6}$$

该算法的信号重建过程是分解过程的逆过程，在图 4 中表示为从右往左逆运算即可得到重构信号，可用式(7)表示。

$$x_n^{(j-1)}=\sum_k h(n-2k)x_k^{(j)}+\sum_k g(n-2k)d_k^j \tag{7}$$

3.1.2　基于 Mallat 的强制阈值除噪算法

在 Mallat 算法的基础上引入强制阈值的思想形成除噪算法，如下所述。

输入：网络信号 x(n)

输出：去噪后的信号 y

步骤 1　用 Mallat 算法对网络信号进行 j 层小波分解；

步骤 2　将分解后的 1 到 j 层细节信号全部置 0；

步骤 3　应用 Mallat 算法重构第(j−1)层的逼近信号，结果为信号 y。

结合之前描述的 Mallat 算法，上述步骤 3 可以用式(8)计算。

$$y=x_k^{(j-1)}=\sum_k h(n-2k)x_k^{(j)} \tag{8}$$

图 6 是原信号 x，可以看到其噪声较多，波形非常不平滑，看不出信号的趋势。通过上述强制阈值去噪算法过滤后得到的信号 y 如图 7 所示。计算过程中采用 coif2 小波，进行 6 层分解，可以看出滤除了高频部分的信号的波形比原信号平滑了很多，可以明显看出信号的走势，说明基于 Mallat 的强制阈值除噪算法是有效的。

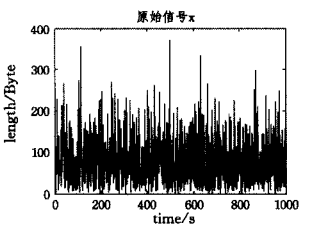

图 6　原信号 x

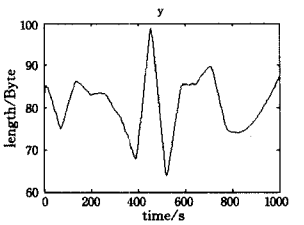

图 7　去噪后的重构信号 y

3.2　基于包速率的详细心跳行为识别算法

用基于 Mallat 算法的强制阈值除噪算法处理原信号得到的近似心跳行为信号，虽然可以看出大致的周期性，但并不是准确的心跳信号。为使判断准确并且求出心跳行为的详细信息，提出基于包速率的详细心跳行为识别算法。

定义 4（数据包的类集合）　在一定时间 T 内与数据包 a_k 有相同的源 IP 地址、目的 IP 地址、负载的数据包的集合称作数据包 a_k 的类集合，记作 $C(a_k)$。

定义 5（数据包 a_k 的速率）　对于数据包 a_k 在时间段 T 内的类集合 $C(a_k)$，速率 G 定义为集合 $C(a_k)$ 内元素的个数与时间 T 的比值，如式(9)所示：

$$G(a_k)=\frac{|C(a_k)|}{T} \tag{9}$$

传统木马的心跳数据包到达的周期较长，并且数据包负载小。因此，为了提高检测效率，需在信号 a 上设阈值 F，以筛选出待分析的数据包。然后利用包速率值判断是否是心

跳包，最后通过计算数据包到达的周期确定心跳行为的详细信息。算法如下所述。

输入：近似心跳信号 y

输出：判断结果

步骤 1　求得近似心跳信号 y 上值满足阈值 F 的数据包集合 B；

步骤 2　求数据包集合 B 中每个元素的类集合，并计算包速率，得到包速率集合；

步骤 3　若存在包速率 G 在阈值 T 内的类集合，则初步认为是心跳包，转步骤 4，否则转步骤 5；

步骤 4　判断包速率集合中的每一个元素是否满足阈值 T，将满足阈值 T 的数据包存下来作为集合 C，若 C 不为空，则存在心跳行为，转步骤 5；若集合 C 为空集，则无心跳行为，转步骤 6；

步骤 5　对 C 集合求心跳频率；

步骤 6　输出判断结果：无心跳行为。

4　实验分析

本文实验分别在有木马的网络环境中和无木马的网络环境中进行检测。实验环境为 30 台台式计算机和交换机组成的小型局域网。计算机 A 的 CPU 为 Intel Core(TM) i5-2400，内存：2GB，操作系统 Windows XP。计算机 B 的 CPU 为 Intel Core(TM) i5-2400，内存：4GB，操作系统：Window7。

实验目的是检测计算机 A 上是否有木马心跳行为。有木马的网络环境即在计算机 A 上安装 Windows XP 虚拟机运行木马的服务端，并在此机器上安装数据包捕获工具，计算机 B 上运行木马的控制端。实验过程中在计算机 B 上对木马进行一些操作以增加干扰数据包。实验的具体步骤如下所示。

(1)使用数据包捕获工具采集网络中的 Tcp 数据包，采集时间为 10 分钟。然后滤掉 IP 数据包长度大于 90 的数据包，形成数据包长度的时间序列 $y=\{y_1, y_2, \cdots, y_m\}$，设数据包 y_i 到达的时刻为 $t(y_i)$，单位为秒。

(2)处理包长度序列 y，形成包长度信号 $x(n)$。由于数据包到达的时刻是不均匀的，不能直接将包长度序列 y 看作包长度信号 $x(n)$，需要在数据包间适当地插入空点。这里处理的算法如下：

$t(y_k)-t(y_{k-1})\geqslant 1$
for k=2:n
　if $t(y_k)-t(y_{k-1})\geqslant 1$
　在第 k 个数据包与 k−1 个包之间插入$[t(y_k)-t(y_{k-1})]\times 10$ 个包长度为 100 的点；
　　end
　end

(3)用 Matlab 实现的基于 Mallat 的强制阈值除噪算法对包长度信号 $x(n)$ 进行处理，得到信号 y。计算过程采用 coif2 小波，进行 6 层分解。

(4)对信号 y 运用详细心跳行为识别算法进行计算。其中阈值 F 设为(0，10]，包速率阈值 T 设为[3/10min，100/10min]。

表 1 是在局域网中运行程序的心跳行为检测结果。实验过程中，虽然对木马程序进行了一系列操作从而造成了干扰，但仍然检测出了心跳行为，该方法能够避免木马其他操作行为对心跳行为造成的干扰；并且运行 SecureCRT 程序，用该方法进行检测，没检测到心跳信号，说明该方法不易误报。除表中的结果外，又进行了大量实验，在局域网中检测的正确率可达 96%，比现有的大部分算法的准确率都高。但是，在互联网环境下本算法的检测率会降低到 85%，容易受到网络流量上传下载的影响，该方法还有待进一步提高。

表 1　部分程序实验结果

计算机 A 上运行的程序名称	心跳频率	计算机 A 上运行的程序名称	心跳频率
流萤	30s	灰鸽子	30s
任我行	30s	PCshare	120s
Finalfantasy	10s	biforst	15s
Ghost	120s	SecureCRT	无心跳

结束语　传统的木马心跳行为检测方法会受到木马自身其他行为数据包的干扰。为解决此问题，本文将信号处理的思想应用到心跳行为检测中，提出基于小波分析的木马心跳检测方法。首先定义了包长度信号来描述网络数据包流；然后设计了基于 Mallat 的强制阈值除噪算法本处理信号得到近似心跳信号，以此过滤掉许多无关的数据包，能够增强抗干扰性；最后提出行为详细信息识别算法，通过此算法的计算能够得出检测结果。实验表明，该方法能够有效地检测出木马心跳行为并且不受木马其他传输数据包的影响。但是该方法在互联网中的检测效果仍不理想，有待优化。

参 考 文 献

[1] Scarfone K, Mell P. Guide to intrusion detection and prevention systems (idps)[J]. National Institute of Standards and Technology Special Publication, 2007, 2007(800): 94

[2] Ding Wei-qiu, A detection technology based on network behavioral characteristics of Trojan [D]. Nanjing: Nanjing University of Posts and Telecommunications, 2013(in Chinese)
丁卫球. 基于网络行为特征的木马检测技术[D]. 南京：南京邮电大学，2013

[3] Xia Ai-min, Zhang Hong-zhi, Yang Wei-feng. Trojan Horse Detection Technology based on Characteristics of Comprehensive Behavior[J]. Information Security and Communications Privacy, 2014(6): 109-113(in Chinese)
夏爱民，张宏志，杨伟锋. 基于综合行为特征的木马检测技术研究[J]. 信息安全与通信保密，2014(6)：109-113

[4] Ma Li-jun. The survey of theft Trojan Detection based on behavior detection[J]. Journal of Guangxi University for Nationalities(Natural Science Edition), 2014, 20(2): 70-74(in Chinese)
马立军. 基于行为检测的窃密型木马检测研究[J]. 广西民族大学学报(自然科学版)，2014，20(2)：70-74

[5] Tao He, Hao Zhong. Network heartbeat packets recognition based on DTW and HC-FCM algorithm[C]// 2010 Sixth International Conference on Natural Computation (ICNC). IEEE, 2010, 6: 3190-3193

[6] Yi Jun-kai, Chen Li, Sun Jian-wei. Data flow clustering detection approach of network heartbeat packet sequence[J]. Computer Engineering, 2011, 37(24): 61-63

[7] Pu Yi-guo, Chen Xiao-jun, Cui Xu, et al. Data Stolen Trojan Detection based on Network Behaviors[J]. Procedia Computer Science, 2013, 17: 828-835

[8] Meng Lei, Liu Sheng-li, Liu Long, et al. Trojan Rapid Detection Method Based on Heartbeat Behavior Analysis[J]. Computer Engineering 2012, 38(14): 13-16(in Chinese)
孟磊，刘胜利，刘龙，等. 基于心跳行为分析的木马快速检测方法[J]. 计算机工程，2012，38(14)：13-16

[9] Wu Xiao-pei, Song Jun-ke, Guo Xiao-jing, et al. The Online En-

velope Detection Based on Sliding Window ICA and Its Application to Brain-Computer Interface[J]. Acta Biophysica Sinica, 2012,28(11):896-909(in Chinese)
吴小培,宋俊可,郭晓静,等. 基于滑动窗独立分量分析的在线包络检测新方法及其在脑-机接口中的应用[J]. 生物物理学报, 2012,28(11):896-909

[10] Oppenheim A V, Willsky A S, Nawab S H. Signals and systems [M]. Englewood Cliffs, NJ: Prentice-Hall, 1983

[11] Yang Yue-xiang. The Research on the Algorithms of Information Hiding and Network Traffic Detection Based on Wavelet Analysis[D]. Changsha: National University of Defense Technology, 2008(in Chinese)
杨岳湘. 基于小波变换的信息隐藏与网络流量检测方法研究[D]. 长沙:国防科学技术大学,2008

[12] Yang Ji-peng, Liu Xue-cheng. Study of The NetWork Abnormal Detection Based on The Wavelet Transforms[J]. Journal of Shanghai Agicultural University(Natural Science), 2011(1):95-99(in Chinese)
杨继鹏,刘学诚. 基于小波变换的网络异常检测研究[J]. 山东农业大学学报(自然科学版),2011(1):95-99

(上接第 121 页)

由规则 4 可知,$p' \in_{t'} (PT \cup PD)$,因为受控进程的状态不可能转换为非可信进程和可信进程,与 $p \rightarrow_{t'} f'$ 类似递归分析 $p \rightarrow_{t'} f'$,必然存在 p^0 在 t_0 时刻直接读取 f,且 $p^0 \in_{t_0} (PT \cup PD)$,由(1)知,也存在矛盾。

综上所述,假设 $\exists f \in_{t_0} DS$, $\exists t > t_0$, $f' \in_t DL$, $f \rightarrow_t f'$ 不合理,故结论成立。

根据定理 1 可知,云盘内文件内容在访问和使用过程中将被限制于云盘和临时安全缓存内;由于云盘和临时安全缓存内的文件数据是加密存储,且只能由受控进程访问,因此外部威胁主体(如木马、病毒)进程无法解密访问,而内部威胁主体即企业内部成员,虽然通过受控进程可以解密访问,但是无法携带明文数据离开该主机和云盘系统(除非在信息安全管理员授权许可的情况下),避免泄密事件发生。

结束语 根据云盘的安全需求,本文提出了一种基于虚拟隔离机制的云盘安全访问模型。该模型主要由安全云盘、虚拟隔离运行环境和 I/O 代理进程组成,能够有效保证数据的受控访问并防止泄漏。其中,虚拟隔离运行环境 VIEE 是该模型的核心,本文重点描述了 VIEE 的设计与实现。VIEE 把可信网络、临时安全存储、云盘和受控进程划为一个安全域,利用网络访问过滤技术、内存空间保护技术和文件过滤驱动技术构建虚拟隔离运行环境,使数据只会流动于安全域内,从而防止敏感数据外泄。随后,基于虚拟隔离机制的云盘安全访问模型,本文实现了原型系统(CFS),测试并分析了终端主机上 CFS 系统对文件读写操作性能的影响。最后,通过安全性分析,证明了基于虚拟隔离机制的云盘安全访问模型的数据安全性。

在下一步研究中,将进一步测试该模型对各种应用软件的性能影响,分析性能影响由哪些因素导致,从而提出算法优化方案。

参 考 文 献

[1] https://www.dropbox.com

[2] https://skydrive.live.com

[3] http://www.ksyun.com

[4] http://yunpan.360.cn

[5] Cao Xi, Xu Li, Chen Lan-xiang. Data integrity verification protocol in cloud storage system[J]. Journal of Computer Applications, 2012,32(1):8-12(in Chinese)
曹夕,许力,陈兰香. 云存储系统中数据完整性验证协议[J]. 计算机应用,2012,32(1):8-12

[6] Yan Xiang-tao, Li Yi-fa. Integrity Checking Algorithm Based on Hash Tree for Cloud Storage[J]. Computer Science, 2012, 39(12):94-97(in Chinese)
颜湘涛,李益发. 基于哈希树的云存储完整性检测算法[J]. 计算机科学,2012,39(12):94-97

[7] Liu Fan, Yang Ming. Ciphertext policy attribute based on encryption scheme for cloud storage[J]. Application Research of Computers, 2012,29(4):1452-1456(in Chinese)
刘帆,杨明. 一种用于云存储的密文策略属性基加密方案[J]. 计算机应用研究,2012,29(4):1452-1456

[8] Lim H, Kapoor V, Wighe C. Active disk file system: Adistributed, scalable filesystem[C]// Proceedeths of the Eighteenth IEEE Symposium. Washington, DC: IEEE Computer Society, 2001:101-114

[9] Swank J D, Goodson G R, Scheinholtz M L, et al. Self-securing storage: Protecting data incompromised systems[C]// Proc of the 4th Symposium on Operating Systems Design and Implementation. Berkeley, CA: USENIX Association, 2000:12-26

[10] Jin Chao, Zhen Wei-min, Zhang You-hui. Active Storage Architecture[J]. Chinese Journal of Computers, 2005, 28(6): 1013-1020(in Chinese)
靳超,郑纬民,张悠慧. 主动存储系统结构[J]. 计算机学报, 2005,28(6):1013-1020

[11] Zhao Yue-long, Jiang Qian. Research and Design of the Virtualization Storage Technology Based on Intelligent Network Disk [J]. Journal of Computer Research and Development, 2009, 46(Suppl):44-49(in Chinese)
赵跃龙,蒋骞. 基于智能网络磁盘的虚拟存储技术的研究与设计[J]. 计算机研究与发展,2009,46(Suppl):44-49

[12] Using virtual desktop to improve the ability of the management and control [EB/OL]. [2010-3-15]. http://www.vmware.com/cn/solutions/desktop

[13] King S T, Dunlap G W, Chen P M. Operaring System Support for Virtual Machines[C]// Proc of the 2003 Annual USENIX Technical Conference. 2003:6

[14] Griffin J L, Jaeger T, Perez R, et al. Trusted virtual domains: Toward secure distributed services[C]// Proceedings of the 1st IEEE Workshop on Hot Topics in System Dependability. Los Alamitos: IEEE Computer Society. 2005:274-281

[15] Gasmi Y, Sadeghi A-R, Stewin P, et al. Flexible and Secure Enterprise Rights Management Based on Trusted Virtual Domains [C]// Proc of the 3rd ACM Workshop on Scalable Trusted Computing. 2008:71-88

[16] Yang Yu. OS level virtualization and its applications [D]. New York: Stony Brook University, 2007

[17] http://dokan-dev.net/en

[18] Denning D E. A lattice model of secure information flow[J]. Communications of the ACM, 1976, 19(5):236-243