基于进程防火墙与虚拟盘的非法信息流过滤方法*

张雪峰,周顺先
(广州番禺职业技术学院,广东 广州 511483)

摘 要:提出一种在虚拟磁盘中对文件映像前后的访问进程进行监控并对非法信息流进行过滤的方法。该方法在关键字或特征信息提取过程中安装各种钩子并对接入主机进行审计,用来提高系统服务器的包转发速率与非法信息的捕捉能力。其目的是阻止病毒或木马程序对文件破坏或数据包劫持,保证信息接入的可控性和安全性。实验测试表明,系统降低了包转发时延,提高了包转发速率和非法信息的识别能力。

关键词:网络信息安全;进程防火墙;虚拟磁盘技术;非法信息流;内核钩子;过滤代价
中图分类号:TP309 **文献标识码**:A **文章编号**:1674-7720(2013)20-0051-03

A model of illegal information flow filtering based on process firewall and virtual disk technology

Zhang Xuefeng, Zhou Shunxian
(Guangzhou Panyu Polytechnic, Guangzhou 511483, China)

Abstract: The paper presents a method that monitors the access process of the virtual disk file image before and after and using illegal information flow filter drivers. There install a lot of host audit hooks on host audit system during the extraction of the keywords or characteristic information, in order to improve the packet forwarding rate of the server and the capturing ability of illegal information. Its purpose is to prevent document from destruction by virus or Trojan program, and achieve controllability and safety in information access. Finally, we perform an experiment through IxLoad. The results of the experiment show that the system can reduce the packet forward delay, improve the packet forwarding rate and better the recognizing ability of illegal information.

Key words: network information security; process firewall; virtual disk technology; illegal information flow; kernel hooks; filtering cost

目前,虚拟磁盘技术广泛应用于快速安装各类(如ISO、BIN等)光盘映像文件中,将文件映射为一个虚拟磁盘,以实现用户的透明性文件操作。非法信息流过滤驱动通常有两种实现方法:一种是以微软的Sfilter模型为基础、以过滤底层文件操作IRP包为技术的传统模型[1],另一种是以微软的MiniFilter模型为基础、以过滤底层文件操作事件为技术的新式模型[2]。

在以包过滤为基础的防火墙中,王洁实现了一种基于FPGA的网络硬件防火墙,以内容过滤设计准则实现对数据包的处理逻辑,使系统免受硬件恶意后门和软件安全漏洞的影响,其劣势在于处理速度远低于目前专用处理器的速度[3]。赵跃华等利用专家知识检测网络层数据包的攻击行为和运行中应用程序的攻击行为推理,可实现防火墙过滤规则的动态生成[4]。侯整风在参考文献[5]中讨论了多核防火墙分层内容过滤的时延问题。温贵江采用了Winsock2套接字技术,在Socket中插入一层,完成传输质量控制、扩展TCP/IP协议栈、URL过滤及网络安全控制等功能[6]。可以看出,这些文献主要涉及内容过滤、规则推理等方法对入侵信息进行识别与阻止,较少考虑内核审计与实时监控。

本文通过在虚拟磁盘中对文件映像前后的访问进程进行监控并对非法信息流进行监控、查找、定位的过滤方法,能对虚拟磁盘的原映像文件进行访问,可识别隐藏在底层的非法信息流,并实现分组过滤与实时监控,用来阻止病毒或木马程序对传输信息的破坏或数据包的劫持,维护了正常的主机接入、包访问的可控性和数据流信息的安全性。

*基金项目:国家自然科学基金资助项目(51074097)

1 模型系统的构成

1.1 系统结构

为实现对内核底层分组识别过滤与监控，系统在通用操作系统的基础上增加应用层的进程策略管理部分和内核层非法信息流过滤驱动，其结构由应用层管理程序、非法信息流过滤驱动程序及虚拟磁盘驱动3部分组成。其中非法信息流过滤驱动负责对虚拟磁盘的原映像文件及映像后的虚拟盘文件的各种操作（包括文件列表、文件数据读写、文件属性读写等）并进行过滤，在过滤例程中对访问进程进行控制，可配置成只有本系统的应用层管理程序能对虚拟磁盘的原映像文件进行访问，只有桌面程序及其他文件关联的应用程序(如OFFICE程序等)能对映像后的虚拟盘文件进行访问，其信息流过滤系统结构如图1所示。

图1 信息流过滤系统结构

1.2 非法信息流过滤驱动的设计

(1)非法信息流的确定

先搜索目标关键字(或依据信息流特征捕捉)，后搜索要剔除的关键字，可降低误判合法信息流的几率。过滤驱动的非法信息流过滤过程如图2所示，其识别与过滤步骤为：(1)在页面中搜索并确定目标关键字(或信息流特征)；(2)标记目标关键字的位置、所在段落编码、字长、数据包头信息等；(3)在标记段落中寻找要剔除的关键字，如果找到存在非法信息，则判定该页面不是要检查的目的页面；如果不存在，则继续搜索该页面的后续内容；(4)将合法信息流关键字(或信息流特征)映像虚拟磁盘，完成过滤驱动的非法信息流过滤过程。

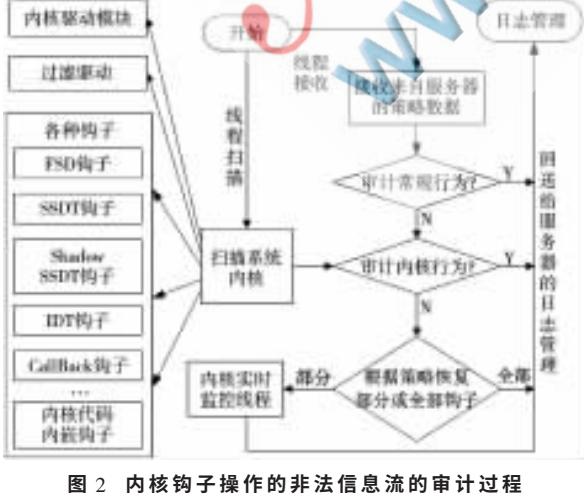

图2 内核钩子操作的非法信息流的审计过程

(2)驱动程序的设计

程序从IRP头部所指定的内存中读取用户数据，如果是Direct方式，则应从IRP头部的MdlAddress所描述的内存中读取用户数据，最后将数据写入到硬件设备中。如果用METHOD_BUFFERED方式，过滤驱动的输入、输出缓冲区都由AssociatedIrp.SystemBuffer指定。

2 非法信息流过滤过程的实施

非法信息流过滤过程主要涉及关键字(或特征信息)的识别与分离、特征筛选、非法信息认定、映像并缓存磁盘。

过滤非法信息的方法在参考文献[7-9]中都有描述，本文通过Windows内核反钩子技术与内核钩子还原技术来识别与分离非法信息，尽量防止信息伪装并通过防火墙来连接主机。

2.1 关键字或特征信息的抽取

关键字或特征信息的抽取方式通过安装各种钩子审计实现，钩子的种类很多，每种钩子可以截获并处理相应的消息，如键盘钩子可以截获键盘消息，鼠标钩子可以截获鼠标消息，外壳钩子可以截获启动和关闭应用程序的消息，日志钩子可以监视和记录输入事件。采用内核反钩子技术的主机审计与传统的主机审计最大的不同在于多了内核钩子扫描模块，并能根据服务器预先设定的策略来对内核钩子进行操作，其对内核钩子操作的审计流程如图2所示。

防火墙内核钩子操作的审计流程图说明从审计内核行为未通过审计操作的行为中，可确定为非法信息流，并记录下线程在段落位置，根据策略恢复内核局部或全部钩子并回送给服务器的日志管理数据库缓存。

2.2 信息流过滤及代价

首先，设定的策略为多组非法的行为信息，每组行为信息中包含至少一个安装NDIS网络通信钩子行为。其次，将内核钩子的行为与设定的每组行为信息进行比较，如果当前内核钩子的行为与某一组行为信息相匹配，则确定执行该组行为信息的钩子为非法，从而过滤相应非法信息流。

过滤代价α的大小主要涉及网络带宽β_1、CPU资源占用比β_2、缓存β_3、审计时间β_4、转发时延β_5、连接速度β_6及误判率β_7等因素集合。

设环境条件为集合$\gamma=\{\beta_1,\beta_2,\beta_3,\beta_4,\beta_5,\beta_6,\beta_7,\cdots\}$，其过滤代价$(\alpha|\gamma)$为$\{\alpha|\beta_1,\beta_2,\beta_3,\beta_4,\beta_5,\beta_6,\beta_7,\cdots\}$，以本文提到的基于进程防火墙与虚拟磁盘技术的非法信息流过滤模型为例，比较在不同环境条件下的过滤代价，通过IXIA公司的IxLoad软件获得相应β参数，审计时间即进程连接开始到回送给服务器的日志管理间隔。

与采用防火墙与虚拟磁盘技术的非法信息流过滤方法前后的系统性能比较，内核反钩子技术更侧重于控

制系统底层信息的读取及伪装代码的识别过程，不同于主机非法接入阻断[10]。与其他 β 参数比较，网络带宽的控制可以通过路由直接设定，因而，利用不同网络带宽更能反映出采用两种不同方法的过滤效果。

3 性能测试及分析

3.1 测试环境与过程

3.1.1 实验环境

进行非法信息流过滤测试。由 1 套 Secoway USG2220 防火墙及其相应模块、1 套 IPtables、1 台 H3C LS-3100-26TP-SI-H3 网络工作组交换机、1 台 H3C ER6300 小型路由器等，通过修改网卡协议、流程监控程序，共同构成实验测试环境。

3.1.2 实验测试步骤

(1)设定的过滤策略为多组非法的行为信息，每组行为信息中包含至少一个行为(即 β_i)；

(2)测试传统的主机审计对非法信息流过滤的各 β_i 值；

(3)当使用内核钩子扫描模块对某一软件的内核进行扫描时，获得过滤代价($\alpha|\gamma$)的各因素 β_i，重复 50 次实验测试并计算平均值。

3.2 结果分析

从防火墙内核钩子操作的审计流程开始到暂存到日志管理过程，设置系统网络带宽 β_1 不断增加时，监测到服务器各项参数值，取 50 次测试所得的平均值，得到 CPU 资源占用比 $\overline{\beta_2}$、审计时间 $\overline{\beta_4}$、转发时延 $\overline{\beta_5}$、包转发速率 φ 曲线，即普通个人版防火墙审计过滤代价(曲线 A)与采用基于进程防火墙与虚拟磁盘的内核反钩子技术的非法信息流过滤代价(曲线 B)，如图 3~图 6 所示，系统服务器上网卡速率为100/1 000 Mb/s、PC 上的网卡速率为 100 Mb/s。

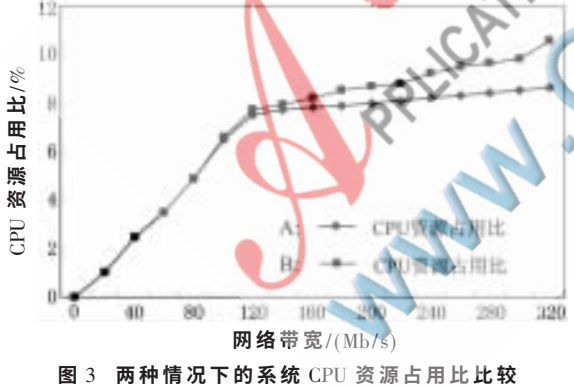

图 3 两种情况下的系统 CPU 资源占用比比较

从图 3 可以看出，系统网络带宽在100 Mb/s 以下两种情况 CPU 资源占用比基本相同，在网络带宽大于 100 Mb/s 时采用基于进程防火墙与虚拟磁盘的内核反钩子技术的 CPU 资源占用比略高于通用防火墙情况，主要因审计内核流程变长、对访问进程进行控制造成，但 CPU 资源占有比仍保持在 11%以下，不会对系统性能造成大的影响。

由于实际系统可能造成一定程度的网络瓶颈。当服务器设定不同网络带宽时，系统过滤代价随之变化，可

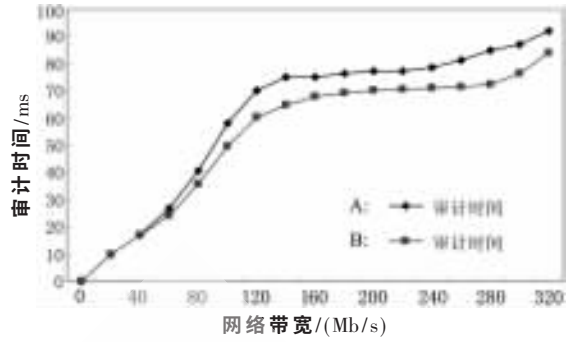

图 4 两种情况下的系统审计时间比较

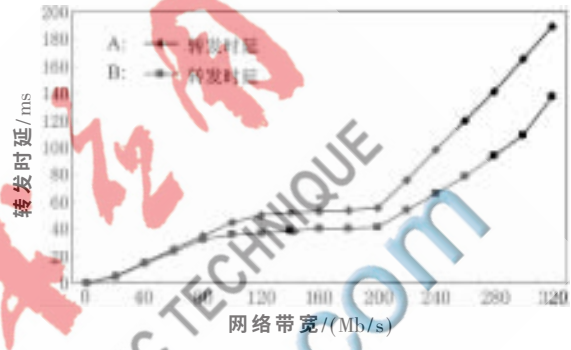

图 5 两种情况下的系统转发时延比较

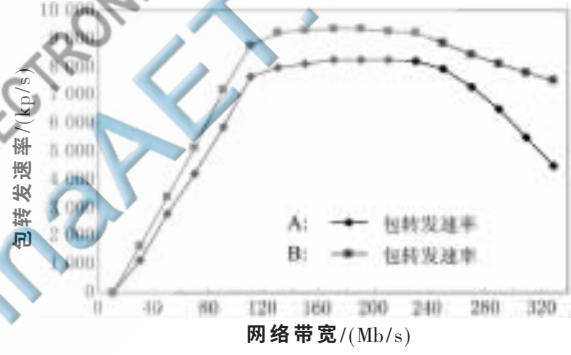

图 6 两种情况下的系统包转发速率比较

以得出：

(1)当 β_1<80 Mb/s 时，系统 CPU 资源占用比、审计时间、转发时延、包转发速率受影响比较小，普通个人版防火墙审计过滤代价(曲线 A)与采用基于进程防火墙与虚拟磁盘的内核反钩子技术的非法信息流过滤代价(曲线 B)基本一致，没有太大的变化，主要是因为带宽较低，系统冗余较大、负荷较少，系统转发时延影响最小(如图 3、图 5 所示)，图 6 反映本文提到的方法系统包转发速率优于普通个人防火墙。

(2)当 80 Mb/s≤β_1≤200 Mb/s 时，图 3~图 6 中曲线变化比较平稳，反映过滤代价 α 在过滤过程中系统较为稳定地运行；图 4~图 6 中，采用本文方法的系统审计时间、包转发速率等性能不仅优于普通防火墙而且性能稳定，采用虚拟磁盘的内核反钩子技术的转发时延降低，包转发速率加快。

(3)当 β_1>200 Mb/s 时，两种情况系统过滤代价等性

能随之降低，随着网络负荷增加，采用内核钩子技术其过滤代价曲线相对平稳、变化趋势较缓和，图4~图6说明此时仍有较好的稳定性。

在本系统环境下，每次接收2万个数据包，进行20次测试，得出平均值，获得在普通个人版防火墙审计过滤情况（非法信息过滤率约99.996%）与采用基于进程防火墙与虚拟磁盘的内核反钩子技术的非法信息流过滤情况（非法信息过滤率约99.998%），在非法信息捕捉能力要好过普通防火墙审计，系统审计时间并没有明显降低，通过合法信息流映射虚拟磁盘技术，降低了转发时延，提高了包转发速率。

在信息安全领域对非法信息审计时采用进程防火墙与虚拟磁盘技术、内核钩子技术能有效改善信息流中非法信息的捕捉能力，其具体表现在：一是采用关键字（或信息流特征）映像虚拟磁盘减少寻道时间，提高包转发速率；二是在合适的网络带宽（例如本系统80 Mb/s≤β_1≤200 Mb/s）时，高效发挥了CPU、缓存和防火墙的作用，实际测试中的审计时间不会大幅增加，但降低了系统转发时延，并提高了系统服务器的包转发速率和非法信息的捕捉能力。

参考文献

[1] ANDREI B，MICHAEL M.Network applications of bloom filters:a survey[J].Internet Mathematics，2005，1(4)：485-509.

[2] FREDRIKSSON K.On-line approximate string matching in natural language[J].Fundamenta Informaticae，2006，72(4)：453-466.

[3] 王洁.基于FPGA的硬件防火墙内容过滤技术研究[D].哈尔滨：哈尔滨工业大学，2009.

[4] 赵跃华，周万胜.防火墙过滤规则动态生成方案设计[J].计算机工程，2012(2)：135-137，140.

[5] 侯整风，庞有祥.多核防火墙分层内容过滤的时延分析[J].计算机工程与应用，2011(12)：93-96.

[6] 温贵江.基于数据包过滤技术的个人防火墙系统设计与研究[D].吉林：吉林大学，2010.

[7] 胡连勇.基于Netfilter框架的校验字过滤防火墙的设计与实现[D].成都：电子科技大学，2007.

[8] 黄利斌，寇雅楠.基于依存句法的网页内容防火墙设计[J].计算机工程与设计，2011(5)：1597-1560，1608.

[9] 杜飞.基于特征字的病毒过滤防火墙技术研究[D].北京：北方工业大学，2010.

[10] 张雪峰，周顺先.一种基于网络安全设备联动的数据包阻断方法[J].计算机与网络，2011，37(12)：68-71.

（收稿日期：2013-07-29）

作者简介：

张雪峰，男，1976年生，硕士，高级工程师，系统分析师，主要研究方向：网络计算与安全、分布式数据库。

周顺先，男，1968年生，博士，教授，硕士生导师，主要研究方向：计算机网络、机器学习。