文章编号:1008-0570(2006)09-3-0134-03

针对多态病毒的反病毒检测引擎的研究

The Study On Anti-Virus Engine Of Polymorph Virus

(临沂师范学院)王振海　王海峰
Wang,Zhenhai Wang,Haifeng

摘要:计算机病毒严重威胁着计算机系统的安全,多态病毒采用自动变形技术对抗特征码检测,本文介绍了利用虚拟机技术的病毒检测引擎的工作原理,讨论了目前存在的效率问题,提出一个采用启发式扫描的检测引擎模型。

关键词:多态病毒;检测引擎;虚拟机;病毒行为
中图分类号:TP393.08　　**文献标识码:**A

Abstract: Computer systems are being severely threatened by computer viruses. Polymorph virus use the auto polymorph mechanism to avoid being detected by Virus Scanner that identifies the virus characteristic code. This article thoroughly discusses the mechanism of Anti-Virus Engine and finds the problem of the Engine. In the end put forward a new method for Anti-Virus Engine base on heuritic strategy.

Key words:Polymorph virus;Anti-Virus Engine;Virtual Machine;Virus Behavior

技术创新

1 引言

近几年中计算机病毒正以惊人的速度蔓延，对计算机系统的安全构成了威胁。早期的计算机病毒并没有采用自动变形技术,都具有固定的特征码。因此,反病毒软件可以利用病毒特征码匹配很容易的检测出隐藏在系统中的病毒程序。然而,病毒和反病毒技术这种"矛与盾"的斗争中,尽管反病毒专家采用了各种各样的方法来检测计算机病毒，但是新病毒还是层出不穷,而且技术水平越来越高,隐蔽性越来越强。现在许多病毒采用自动变形技术来逃避特征码检测技术的检测,这就是所谓的多态病毒。多态病毒是指每次传染产生的病毒副本在外观形态上都发生变化的病毒。因此,多态病毒在外观形态上没有固定的特征码。

多态病毒之所以能产生自动变形是因为其内部有一种变形机构,本文称之为多态变形引擎(Polymorph Engine)。下面分析多态病毒的变形机制,多态病毒的变形引擎主要由五部分组成: 预处理器(Preprocessor)、还原器(Restorer)、病毒体(Virus Main Code)、变形驱动器(Polymorph Driver)、变形器(Polymorph Processor),五部件相互协作,共同完成传染和变形。预处理器在病毒进入内存时对病毒进行预处理，如将分块寄生的病毒进行组装。还原器在病毒进入内存后将被变形器变形的部分还原。病毒体代码完成普通病毒的常规任务,如传染、破坏等。变形驱动器是对病毒产生变形的变形控制部件。变形驱动器对预处理器和还原器产生代码等价变形,调用变形器对其它部件产生变形。变形器使病毒体代码、变形驱动器和变形器产生变形,还原器是变形器的逆变换器。由于变形器在对同一数据或代码进行两次变形时，所得到的两个结果相同的概率很小，所以假定变形器对同一数据的多次变形运算都会得到不同结果。多态病毒采用的这种程序演化的技术,使基于特征码检测的反病毒软件彻底失去作用。

2 反病毒技术与检测引擎原理

2.1 虚拟机技术

虚拟机技术是一种前沿的反病毒新技术，主要用来分析未知病毒和查、杀多态变形病毒。具体的思想是用程序代码虚拟 CPU、各个寄存器甚至是硬件端口,将采集到的病毒样本放该虚拟环境中执行,通过分析内存和寄存器以及端口的变化来了解程序的执行情况。当虚拟机技术加入病毒检测引擎中,由于该技术采用动态分析程序的变化，对于多态变形病毒和未知病毒的发现准确性很高。因为对于多态病毒,无论如何变化代码和加密代码,可是最终执行时刻还是要现出真面目。但是该技术虚拟的 CPU 执行速度比真正的 CPU 慢 10 多倍,所以在查、杀效率上有待于提高。

2.2 反病毒检测引擎工作原理

分析目前反病毒检测引擎的工作原理，如图 1 所示。

首先文件类型检测模块将检测的目标文件分为二进制可执行文件与文本类型两大类。目前基于文本格式的计算机脚本病毒已经成为目前的主流形

王振海:讲师 硕士

式,文本类型的检测对象如果是 OFFICE 文件,因为其中含有 basic 宏代码,先经过一个预处理器提取其宏代码后交给下级语法分析器处理,分析完语法后再由脚本检测引擎从病毒特征码库中提取脚本检测码进行模式匹配,最后将结果送 GUI 通知用户检测结果;对于普通的 perl,php,js,vbs 文本类的源程序文件就直接交语法分析器处理,省去了提取内含 basic 代码的过程。二进制执行文件先检测是否存在某种类型的加密外壳,如果是一个有外壳的执行文件,则转入一个递归的脱壳模块直到脱出真正的执行代码体;接着是进入检测变形病毒的虚拟机执行一段指令,对存在解密指令段的程序进行变形病毒的检测操作,如果不是则跳出虚拟机进入二进制代码的检测模块,然后将检测的结果通过 GUI 通知用户。根据计算机用户的指令,病毒检测的结果送入杀毒模块中,从杀毒规则库中提取相应的规则生成杀毒脚本,随后送入杀毒引擎中执行该脚本,整个检测、清除计算机文件型病毒的过程结束。病毒检测引擎工作机制就是这样一个过程。

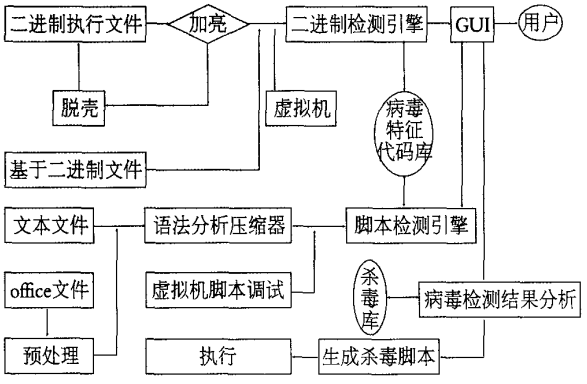

图 1 反病毒检测引擎结构图

2.3 存在的问题

采用虚拟机技术的反病毒检测引擎在实际工作中有很大的效率瓶颈问题,多数检测病毒情况发生在怀疑计算机系统被感染病毒时,此时只有少量的文件被病毒感染,多数文件处于未感染病毒的正常状态,并且由于多态病毒属于编写难度大的病毒,所以多态病毒的数量更少。反病毒软件在使用虚拟机时,是将每个执行文件放入虚拟机中运行一段时间,发现异常后在代码还原的状态下继续使用特征匹配来检测病毒。因此多数情况下,虚拟机的使用是浪费了机器效率来换去查准率。如何解决这个效率瓶颈问题呢?

3 启发式病毒检测引擎

3.1 启发式扫描技术

启发式扫描技术是在软件系统规模趋于庞大,对常用的特征扫描法的扫描速度要求改进的情况下提出的优化特征扫描法的技术。启发式指的"自我发现的能力"或"运用某种方式或方法去判定事物的知识和技能",某种意义上启发式扫描是基于专家系统的原理产生的。由于病毒程序和正常程序在执行行为上的不同,作为汇编级的代码分析人员可以很容易的分析出这类非正常的程序。例如 对正常 PE 文件最后一个节不是执行代码节,而病毒一般把自己添加到正常文件的最后一节,并把执行的入口跳到最后一节。启发式扫描发现这些代码异常之后再对文件进行特征代码扫描,会明显的提高扫描效率。总之,启发式扫描技术的思想是依据一定的先验知识来减小解的搜索空间,提高解的搜索效率。

3.2 病毒行为分析

病毒的行为可以作为启发式扫描的先验知识,病毒程序不同于普通的计算机程序,带有破坏性与复制自身的特征。给每个病毒程序的典型行为分类并说明,组成病毒典型行为特征码数据库,以下给出一种具体分类方案:

(1)D:解密模板,变形病毒必然具有的行为特征。由于变形病毒感染文件时被随机加密,并且在其执行时解密模块是多变的代码,但是病毒设计者是根据一个固定的解密结构利用相同功能程序演化的手段进行随机变化,其解密模板一般是固定的,可以提取其特征码作为重点行为怀疑特征;

(2)G:解密库,因为变形病毒利用解密模板随机生成解密指令,所以必然带有自己的相同功能指令集合的变形指令库。这个指令库具有一定规模并且有很明显的特点,可以作为行为识别特征;(3)F:异常的文件访问,病毒程序在感染时一般要遍历系统中所有的执行文件,这是普通程序一般没有的操作,可以作为重要的行为怀疑特征;(4)A:异常的文件结构,比如 PE 文件头部出现异常标记,这可能是病毒判断感染的标志;PE 文件的最后一个节是可执行属性,这可能就是被病毒感染后添加的病毒体;PE 文件的入口点发生改变等;这些都可以作为行为怀疑特征;(5)M:针对内存区域操作指令数量,病毒在感染和执行时会有大量的内存区域的清除、移动、替换等操作,这类指令可以作为行为怀疑特征;(6)C:修改计算机系统基本配置的指令,比如在注册表中添加启动项、注册服务进程、修改配置文件,由于普通的软件也有这类指令,所以只能作为行为怀疑特征;(7)R:重定位,寄存型病毒程序在其宿主程序中必须进行变量的重定位,这是普通程序所不具有的特点,因此可以作为行为识别特征;(8)! :可疑指令,比如有的病毒运用抗虚拟机分析的指令、为了引起结构异常故意使用的错误指令、无效

跳转指令甚至是 Intel 未公开的指令,这些可作为行为怀疑特征。

3.3 启发式病毒检测引擎

根据图 1(文件类反病毒检测引擎的概念图)发现检测过程中会对大量未被病毒感染的文件进行病毒特征代码的一一匹配,所以通常情况下病毒检测引擎浪费了大量的计算机效率来降低病毒检测的漏报率。在此应用基于病毒行为分析的启发式检测策略。首先,根据病毒程序的行为特征,提取相应的特征字段并对每个行为特征增加一个模糊地权值,比如对于 M 特征(针对内存区域操作的指令数目)设定 3-5 为正常,5-10 为异常,大于 10 条为严重异常。因此,病毒行为特征库设计的首要任务是对各种行为特征模糊量化处理,这项工作可以根据病毒分析专家的经验或是借助人工智能的技术辅助处理。第二,注意分析每个行为特征的关联,根据专家的分析设计简单的权植推理规则,比如 C 特征(修改计算机基本配置)如果只是单独出现此行为,则认为属于正常,因为这可能是某软件的安装程序;如果 C 与 A 特征同时出现,则该文件被病毒感染的概率就提高了,因此其检测权值相应地要提高。例如:如下所示的规则。

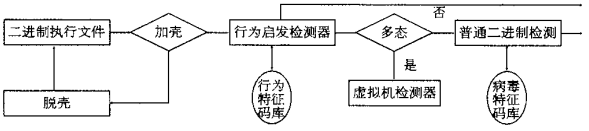

图 2 启发式病毒检测引擎局部结构图

```
Rule1:
If (C == something and other == 0) then
Filecheck.Weight = 0;
Rule2:
If( C == something and A == something) then
Filecheck.Weight=(C.Weight+A.Weight)*2;
```

这种考虑行为特征关联而动态调整文件检测权值的方法有利于提高综合分析病毒程序行为的准确性。在改进的病毒检测引擎中,病毒行为分析器根据简单的规则,分析文件的检测权值,如果超出一定的阈值就送虚拟机处理,作为疑似多态病毒处理;反之,认为是健康文件,放弃该文件的检测工作,因此节省了大量的病毒特征码匹配时间。这样不仅改善了运用虚拟机分析、检测多态病毒的效率瓶颈,而且根据病毒行为特征检测可以预警未知的计算机病毒。

4 总结

本文根据实际情况来考虑针对多态病毒检测引擎的设计,改善了虚拟机技术产生的效率瓶颈,可见对反病毒引擎的设计不能只注重检测算法的改进。本文分析了病毒的行为特性并且依据各关键行为特征为启发搜索的依据提高检测引擎的效率,关键部分是启发式检测思想的应用。另外,有大量的工作需要进一步研究,比如人工智能的机器学习方法来发现合理的行为检测特征;应用统计学与数据挖掘的方法分析准确的检测阈值等。

本文作者创新点:

本文阐述了虚拟机在反病毒应用中的原理,以及分析了多态计算机病毒本质特点。提出了一个基于计算机程序行为的检测方法,该方法根据正常的程序行为和恶意的程序行为来判断是否存在病毒,在行为分析的基础了,运用了模糊处理技术。将程序行为进行模糊量化,不仅提高了反病毒引擎的效率而且降低了误警率。

参考文献:

[1]祝恩,殷建平等.计算机病毒的本质特性分析及检测[J].计算机科学,2001,28(增刊):238-240

[2]唐常杰,胡军.计算机反病毒技术.北京:电子工业出版社,1990-06

[3]宫会丽,丁香乾.GA 和 HS 算法解决电子化配车方法比较[J].微计算机信息,2005,7-3:147-113

[4]Understanding Virus Behavior in 32-bit Operating Environments. Symantec,1997.

[5]Fred Cohen. Computer Virus Theory an Experiments[J].Computer&Security,1987,6(1):22-35

[6]李媛圆,吴灏等.基于免疫原理的可执行文件签名验证模型的研究[J].微计算机信息,2005,12-3:42-45

作者简介:王振海(1972-),男,山东苍山人,临沂师范学院,讲师,硕士.主要研究方向:软件工程,多媒体技术.E-mail:lywzhh@163.com;王海峰(1976-),男,山东临沂人,临沂师范学院,讲师,硕士,主要研究方向:网络安全。

Biography:WANG Zhen-hai,male,born in Jan. 1972,Han nationality,Master,docent,Now he is engaged on teaching and scientific research at Linyi Normal University.His main research fields are software engineering and multimedia technology.

(276005 山东临沂 临沂师范学院 信息学院)王振海 王海峰

(Department of Information Linyi Normal University, Linyi Shandong 276005,China)Wang,Zhenhai Wang,Haifeng

通讯地址:(276005 临沂师范学院 信息学院)王振海

(投稿日期:2006.1.25)(修稿日期:2006.2.27)

书讯

《现场总线技术应用 200 例》
110 元/本(免邮资)汇至

《PLC 应用 200 例》
110 元/本(免邮资)汇至

地址:北京海淀区皂君庙 14 号院鑫雅苑 6 号楼 601 室
微计算机信息杂志收　邮编:100081
电话:010-62132436　　010-62192616(T/F)

万方数据

针对多态病毒的反病毒检测引擎的研究

万方数据 WANFANG DATA 文献链接

作者： 王振海， 王海峰， Wang,Zhenhai， Wang,Haifeng
作者单位： 276005,山东,临沂,临沂师范学院信息学院
刊名： 微计算机信息
英文刊名： CONTROL & MANAGEMENT
年，卷(期)： 2006,22(27)
被引用次数： 10次

参考文献(6条)

1. 祝恩;殷建平 计算机病毒的本质特性分析及检测 2001(zk)

2. 唐常杰;胡军 计算机反病毒技术 1990

3. 宫会丽;丁香乾 GA和HS算法解决电子化配车方法比较[期刊论文]-微计算机信息 2005(21)

4. Understanding Virus Behavior in 32-bit Operating Environments 1997

5. Fred Cohen Computer Virus Theory an Experiments 1987(01)

6. 李媛圆;吴灏 基于免疫原理的可执行文件签名验证模型的研究[期刊论文]-微计算机信息 2005(36)

本文读者也读过(10条)

1. 王振海.王海峰.WANG Zhen-hai.WANG Hai-feng 基于多态病毒行为的启发式扫描检测引擎的研究[期刊论文]-实验室研究与探索2006,25(9)

2. 汪淼.谢余强.舒辉.罗军宏.黄华星 病毒的多态性研究[会议论文]-2004

3. 张秋霞.孙秀英 计算机反病毒引擎策略[期刊论文]-黑龙江科技信息2007(16)

4. 廖翔 网络入侵检测系统中检测引擎的研究与设计[学位论文]2006

5. 崔鹏.CUI Peng 基于语义的启发式病毒检测引擎研究[期刊论文]-常熟理工学院学报2008,22(10)

6. 马宁 应用虚拟机杀毒引擎解决病毒加壳难题[期刊论文]-中国金融电脑2007(1)

7. 李洪敏.凌荣辉 反病毒引擎技术初探[会议论文]-2003

8. 王海峰.夏洪雷.孙冰 基于程序行为特征的病毒检测技术与应用[期刊论文]-计算机系统应用2006(5)

9. 谭云松.Tan Yunsong 一种启发式反病毒技术的研究[期刊论文]-网络安全技术与应用2006(11)

10. 崔鹏.CUI Peng 基于形式化语义的启发式病毒检测引擎研究[期刊论文]-辽东学院学报（自然科学版）2008,15(3)

引证文献(10条)

1. 刘忠民.刘洪.段喜龙 基于用户行为的网络数据过滤方法[期刊论文]-计算机应用与软件 2009(7)

2. 张海燕.肖冬荣.李诗平 计算机病毒入侵及对抗技术[期刊论文]-微计算机信息 2008(9)

3. 王晓洁 蠕虫病毒特征码自动提取原理与设计[期刊论文]-微计算机信息 2007(18)

4. 任师尊 病毒检测技术在查杀"熊猫烧香"中的实证分析[期刊论文]-长春大学学报（自然科学版） 2007(6)

5. 张迎春 基于特征码技术的攻防策略[期刊论文]-计算机系统应用 2009(3)

6. 周晓航.马立玲.王军政 基于数学模型的多进程守护病毒的分析[期刊论文]-微计算机信息 2008(36)

7. 孙海峰.宋丽丽 使用蜜罐分析一种蠕虫病毒的运行机制[期刊论文]-微计算机信息 2008(3)

8. 张永超.张磊.张权.唐朝京 反病毒虚拟机的缺陷及其改进[期刊论文]-信息安全与通信保密 2011(9)

9. 李哲.封汉颖 一类改进的计算机病毒传播模型[期刊论文]-微计算机信息 2008(12)

10. 林永和 校园网防病毒系统的设计和实施[期刊论文]-微计算机信息 2007(12)

本文链接：http://d.g.wanfangdata.com.cn/Periodical_wjsjxx200627048.aspx