二进制病毒的启发式扫描技术

张青霞[1],杨吉峰[2]

(1 聊城大学 农学院,山东 聊城 252000;2聊城大学 实验管理中心, 山东 聊城 252000)

摘　要:传统病毒扫描采用特征值扫描技术,虽然这种方法能够非常有效的查知已知的病毒,但是对新病毒完全无能无力。并且一些加密的多态病毒不可能提取出一个满意的特征值,启发式病毒扫描从一定意义上讲就是一种病毒预测技术,它针对病毒的普遍特征(自我复制,破坏系统),再加上病毒分析师的经验总结,对未知病毒进行有效的预测。

关键字:二进制;宏病毒;虚拟机;特征串

中图分类号:TP309.5　　文献标识码:B　　文章编码:1672-6251(2006)08-0116-02

Dos 时代的病毒大多是由汇编语言写成,那时一个程序非常小,格式简单,易于控制,特别是 com 文件,文件的第一个字节就是程序得以执行的入口, 所以那个时代的病毒非常短小精巧, 甚至有人直接修改其中的机器码,造成了病毒花样倍出,防不胜防。随着高级语言的发展,由高级语言写成的二进制病毒也得以发展,这些病毒的作者再也不能精确控制每一条指令, 因为即使相同的代码,不同的编译器编译的结果也不相同;同时用高级语言编写的病毒另一大特征是库函数的调用,因为它们都不直接使用 DOS 中断和 BIOS 调用了。Widows 平台的不断更新发展造成了可执行文件格式的更加复杂, 人为控制二进制文件的每一个字节几乎不可能, 所以当前病毒作者把更多的精力花在了病毒的传播方式上,而不再追逐代码技巧。

根据不同编译器的固有特点, 经验丰富的程序员能够从文件的二进制格式看出由何种编译器编译而来。由于这种病毒是二进制可执行程序,所以它们几乎可以完全控制整个系统,凡是能想到的破坏方法,它们都能实现。小到拷贝文件,修改文件,系统重启,大到格式化分区,损坏速据,瘫痪网络等。

宏病毒的产生, 是利用了一些数据处理系统内置宏命令编程语言的特性而形成的。这些数据处理系统内置宏编程语言的存在使得宏病毒有机可乘, 病毒可以把特定的宏命令代码附加在指定文件上, 通过文件的打开或关闭来获取控制权, 实现宏命令在不同文件之间的共享和传递, 从而在未经使用者许可的情况下获取某种控制权,达到传染的目的。目前在可被宏病毒感染的系统中, 以微软的 Word、Excel 居多。因为大多数宏病毒是通过感染模板来传播病毒, 所以只要将模板文件的属性改为只读即可。最好将 autoexec.bat 等重要的系统文件也设为只读属性, 因为它们也常常是病毒所攻击的对象。还有一个更简单的办法,就是将自动执行宏的功能禁止掉,这样用户执行任何操作时,病毒都不能依靠自动执行宏的功能来抢占控制权。即使不这样做,也还有办法制止宏病毒的发作:现在的 office 系统在执行宏之前都要询问用户是否执行宏。

使用特征串的扫描法被查病毒软件广泛应用着。当特征串选择得很好时, 病毒检测软件让计算机用户使用起来很方便, 对病毒了解不多的人也能用它来发现病毒。另外,不用专门软件,用 PCTOOLS 等软件也能用特征串扫描法去检测特定病毒。

虚拟机,在反病毒界也被称为通用解密器,已经成为反病毒软件中最引人注目的部分, 尽管反病毒者对于它的运用还远没有达到一个完美的程度, 但虚拟机以其诸如“病毒指令码模拟器”和“Stryker”等多变的名称为反病毒产品的市场销售带来了光明的前景。

结合病毒的两大特性(传染和破坏),在虚拟机的帮助下我们设计如下简单启发式病毒扫描。

1　关注关键指令

不可否认,Intel 的CPU 指令集是不可能被完全模拟,也没有精力对其进行完全理解分析;但更重要的是绝大多数病毒使用的都是常见指令 (从我们分析过的病毒得出的总结),而能够区分病毒和普通程序的指令集合更是如此 (譬如查找待感染文件的 INT 21H 的功能调用)。一般对病毒扫描有用的指令有:文件操作(查找,读写,修改属性,删除)、中断向量表的修改、内存操作(修改,读取,驻留)、常见的赋值,计算,跳转指令、磁盘的直接访问(修改 MBR,分区表)。

2　针对病毒的两大特性分别分析

(1)传染(自我复制)。发现文件操作后,考虑如下

收稿日期:2006-04-19;修回日期:2006-05-22

作者简介:张青霞(1975-),女,助教,研究方向:计算机农业应用。

几种情况:①拷贝指令。找出拷贝的源和目标,“大量”的拷贝当然是值得重点关注;②写指令。找出写的数据来源;③打开,查找文件指令。找出欲打开的文件的特征(比如”*.exe”,”*.com”)。

(2)破坏(对系统的危害)。这个比较复杂,但是它并不是识别病毒的重要依靠,只是一个辅助性的目标,所以我们不要求能找出很多的破坏行为出来,也就是说只关注一些已经出现过的破坏行为,总结如下:①直接修改磁盘;②删除文件;③重启;④格式化硬盘。

3　结果评估

使用的方法就是对不同情况计算权值,然后报告给使用者

目前有两种方法可以跟踪控制病毒的每一步执行。一种是单步和断点跟踪法,和目前一些程序调试器相类似;另一种方法当然就是虚拟执行法。下面分别分析单步和断点跟踪法和虚拟执行法的细节。

单步跟踪和断点是实现传统调试器的最根本技术。单步的工作原理很简单:当 CPU 在执行一条指令之前会先检查标志寄存器,如果发现其中的陷阱标志被设置则会在指令执行结束后引发一个单步陷阱 INT 1H。

用单步和断点跟踪法的唯一一点好处就在于它不用处理每条指令的执行—这意味着它无需编写大量的特定指令处理函数,因为所有的解密代码都交由 CPU 去执行,调试器不过是在代码被单步中断的间隙得到控制权而已。

虚拟执行法的唯一一点缺点就在于它必须在内部处理所有指令的执行—这意味着它需要编写大量的特定指令处理函数来模拟每种指令的执行效果,这里根本不存在何时得到控制权的问题,因为控制权将永远掌握在虚拟机手中。用软件方法模拟 CPU 并非易事,需要对其机制有足够的了解,否则模拟效果将与真实执行相去甚远。但虚拟执行的优点也是很明显的,它正好填补了单步和断点跟踪法所力不能及的方面:首先是不可能被病毒觉察到,因为虚拟机将在其内部缓冲区中为被虚拟执行代码设立专用的堆栈,所以堆栈检查结果与实际执行无二(不会向堆栈中压入单步和断点中断时的返回地址);其次由于虚拟机自身完成指令的解码和地址的计算,所以能够获取每条指令的执行细节并加以控制;最后,最为关键的一条在于虚拟执行确实做到了“虚拟”执行,系统中不会产生代表被执行者的进程,因为被执行者的寄存器组和堆栈等执行要素均在虚拟机内部实现,因而可以认为它在虚拟机地址空间中执行,所以不会有对系统产生破坏的危险。鉴于虚拟执行法诸多的优点,所以将其运用于病毒扫描上是再好不过的了。

4　总结

随着现在病毒复杂性的提高,这种启发式扫描技术我们感觉将有非常广阔的发展前景。在以后的研究中,我将不断致力于二进制病毒的启发式扫描技术的研究工作,为计算机事业的发展贡献自己的力量。

(上接第 135 页)

清水或泥水选种后用恶线清浸种 48~72h,播后晒田,立针后灌浅水,施尿素 100kg/hm² 作断乳肥,移栽前 4~5d 施送嫁肥 75kg 尿素/hm² 左右,并注意防治稻蓟马危害,防治药剂可选用吡虫啉等。

4.2　适时早栽

由于豫南稻区 5、6 月份温度较高,秧苗长得快,秧龄以不超过 25 天为宜,移栽密度 25.5~30 万穴/hm²,要求随拔随栽,最好带水移栽,以减轻秧苗损害,缩短缓苗期。

4.3　肥水管理

大田施 25%复合肥 600kg/hm² 做底肥,移栽后施尿素 150kg/hm² 作分蘖肥,后期根据苗情酌施促花肥、保花肥各 75kg/hm² 左右的尿素。晒田要及时、彻底,当群体达 450 万/hm² 左右时即可晒田,要晒到田面微裂,叶片变黄再复水。由于 III 优 98 分蘖力强,如果晒不好田,容易继续分蘖,造成无效分蘖和小穗增多,影响产量和稻谷质量。收割前晒田宜迟。III 优 98 灌浆慢,灌浆时间长,断水过早容易造成灌浆不饱,影响粒重和稻米质量,在收割前一个星期断水即可。

4.4　病虫害防治

根据近几年的粳稻栽培试验及生产示范的田间观察,该组合没有稻瘟病和白叶枯病发生,纹枯病发生轻。虫害苗期重点防治稻蓟马,中期防治稻纵卷叶螟,抽穗前防治三化螟三代,后期加强对稻飞虱的观测与防治。

参考文献

[1] 张培江,占新春,张明,等.栽插密度和穴栽苗数对优质杂交中粳 III 优 98 产量的影响[J].杂交水稻,2004,19(1):43~44.

[2] 宋世枝,段斌,何世界.豫南粳稻高产优质障碍因子与栽培对策[J].中国农学通报,2005,21(5):206~208.

[3] 张培江.杂交中粳新组合 III 优 98 特征特性及栽培技术[J].安徽农业科学,2003,31(1):71~72.

二进制病毒的启发式扫描技术

作者： 张青霞， 杨吉峰
作者单位： 张青霞(聊城大学,农学院,山东,聊城,252000)， 杨吉峰(聊城大学,实验管理中心,山东,聊城,252000)
刊名： 农业网络信息
英文刊名： AGRICULTURE NETWORK INFORMATION
年，卷(期)： 2006(8)