执行体全量识别与精细管控

安天AVLSDK为执行体治理提供元数据化能力

安天 | 基础引擎中心

目 录

01 为什么要将执行体元数据化

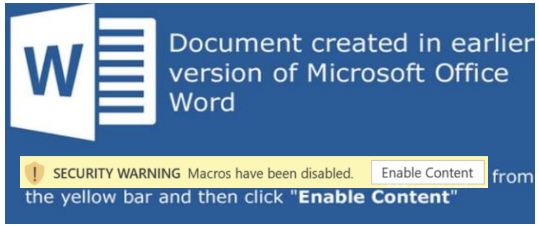
执行体和执行体的元数据化

什么是执行体

- 网络空间的基本组成要素
- 系统运行的基础可观测对象
- 网络治理过程中的基础对象
- 网络空间对抗中的“弹药和装置”

执行体的元数据化

- 执行体的元数据化是指将执行体的元数据和执行体本身分开存储的过程。执行体的元数据包括但不限于**来源、成分、结构、用途、行为、脆弱性**等维度的信息



某APT组织攻击文档

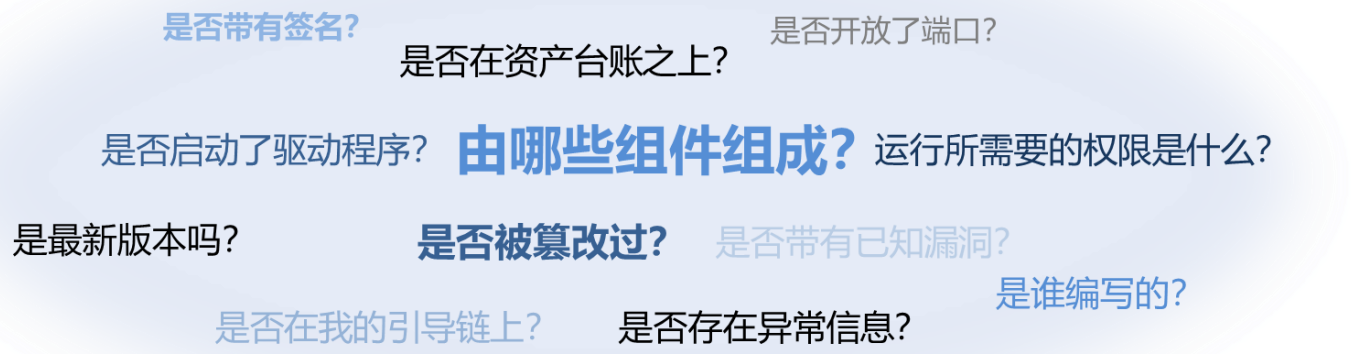
1、来源：作者xxx，编辑时间：2021-06-19
2、文档名称：某军工企业招聘
3、成分信息：内嵌恶意URL[http]://cod.com
4、脆弱信息：cve-2018-11882溢出漏洞
5、异常行为：auto_open宏自启动

为什么要将执行体元数据化

执行体的元数据化是网络治理清晰化的基础，是网空威胁防御的重要支点

粗糙的标识难以支撑精细化的治理

- 粗糙的标识如文件路径、Hash值，很难支撑功能执行逻辑、发布方的判断
- 浅层次的信息不足以支撑批量化的关联、溯源等深层次的工作



执行体元数据化方法

EDR

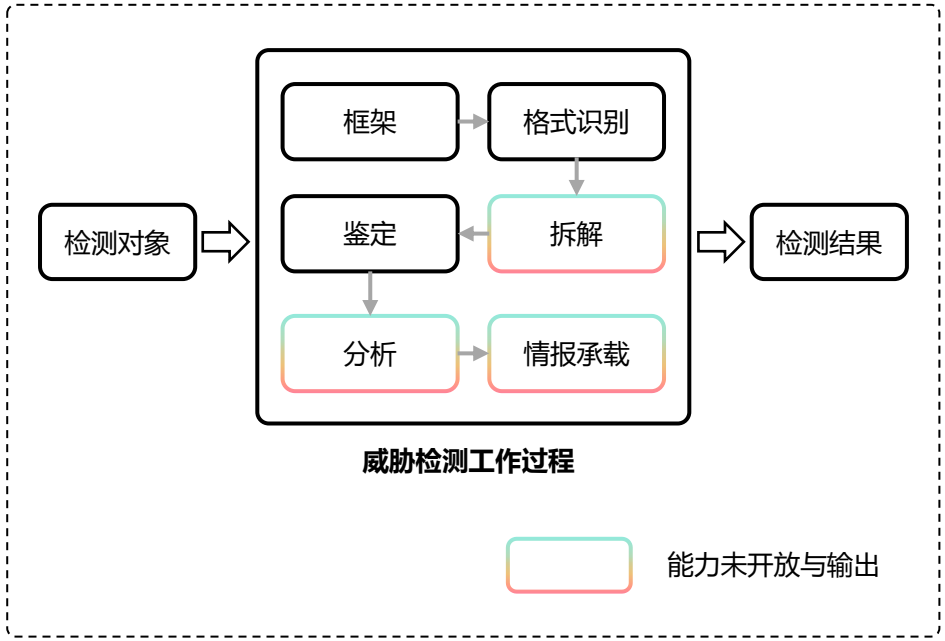
- 可提取文件路径、哈希、大小、IOC等
- 没有提取或充分提取签名、加密混淆手段、组成成分等重要信息

沙箱

- 可揭示细粒度的本地与网络行为
- 但存在执行体无法触发、不能完整执行、效率低的问题

杀毒软件

- 其中间过程静态拆解可提取元数据
- 但主要服务与恶意代码检测，未输出拆解结果



02 安天AVLSDK的执行体元数据化能力

安天AVLSDK可提取的执行体元数据

现有安全软件签名机制的不足

数字签名是可信计算的重要组成部分

- 可以验证软件的真实性并保证它来自合法来源且未被篡改
- 数字证书是数字签名的核心，在诸多层面被广泛的应用，如PE代码签名、固件签名，驱动签名等

现有安全软件签名机制的不足与问题

- **无法在离线场景完成证书的合法性和有效性的验证**

  依赖于联网进行验证，一些拥有被撤销证书的恶意软件可绕过该机制

- **无法检测拥有合法签名的恶意软件**

  通过验证签名是否有效进行运行管控，有的APT组织盗用合法证书给自己的攻击工具签名即可绕过该机制

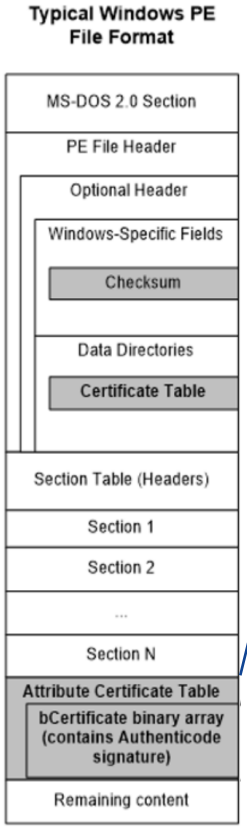
- **无法精准识别身份的信誉度**

  利用身份相关数据构建证书链，而不能判定证书持有者的信誉度。证书持有者签发带有恶意插件（如广告件）的应用，可绕过安全软件签名机制

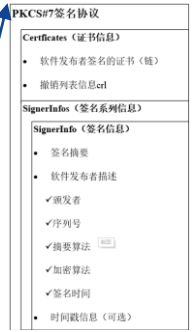
- **无法检测使用假冒签名证书的恶意软件**

  未对证书属性（如签名算法）合法性验证，攻击者利用此攻击点，伪装成可信任的程序，获得签名验证机制对执行体的合法化认证

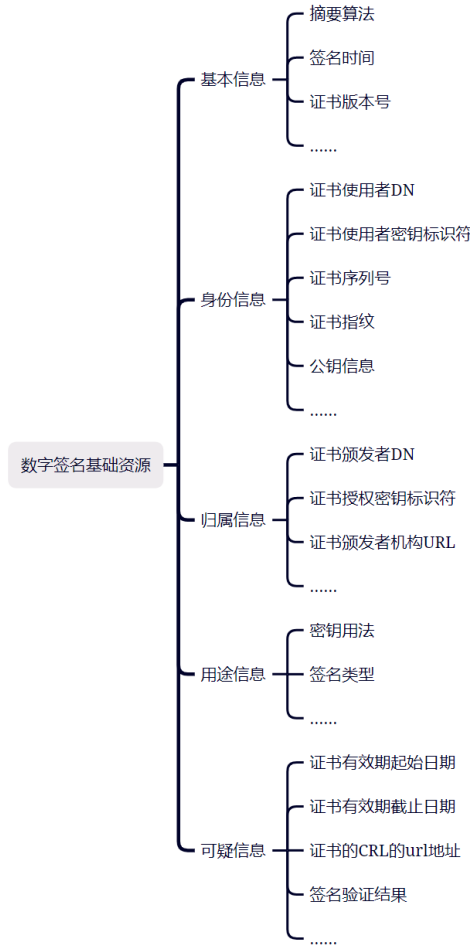
数字签名的元数据化



带有签名的PE文件结构



签名体结构



- **综合判断签名证书的权威性**
- **综合判定证书合法有效性**
- **综合判定签名是否存在密钥泄露等可疑点**
- **综合判定是否被APT组织利用**
- **综合判定是否被广告件利用**

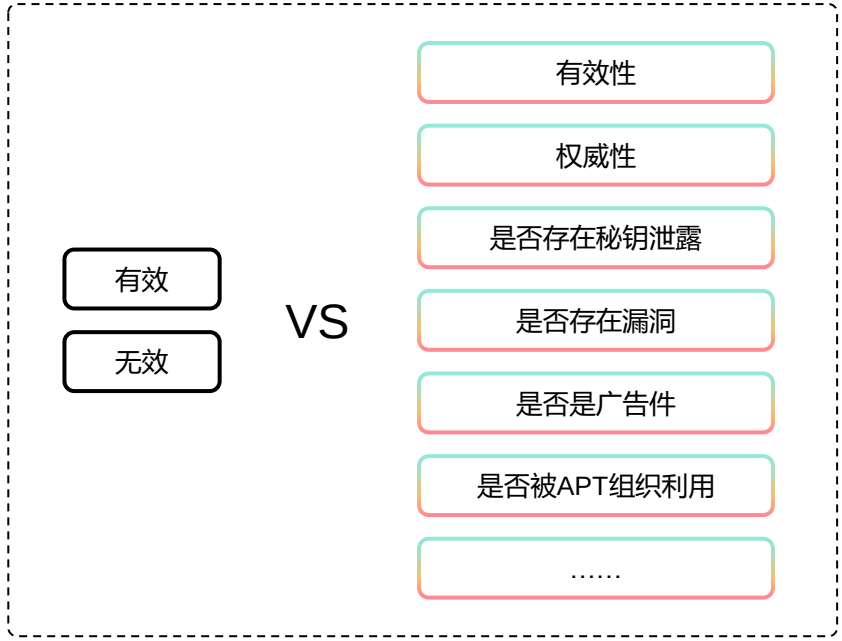
数字签名元数据化的价值

支持用户构建专有化签名清单，形成双信誉机制

- 形成厂商信誉库为辅，用户自定义为主的双重信誉评价机制
- 基于输出的签名元数据，用户可自定义黑白名单或预设相应的守候或管控条件

更收敛更精细的执行体管控

- 通过签名元数据化可形成收敛的签名管理
- 结合HASH维度的治理达成更精细的管控



废止证书离线场景下的验证

场景

- 在离线场景下，因无法连接远程证书查询服务器实时查询证书状态，攻击者使用废止证书也可绕过检测

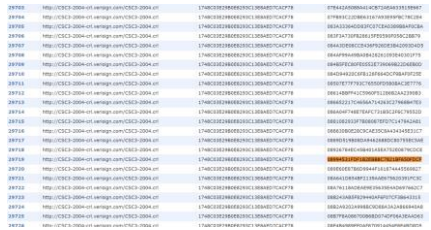
基于元数据化实现离线场景下废止证书的识别

- 安天AVLSDK可对执行体进行身份合法性相关的元数据提取，包括crl下载连接，证书序列号，签发者信息等，结合轻量级离线证书库，综合验证证书合法性

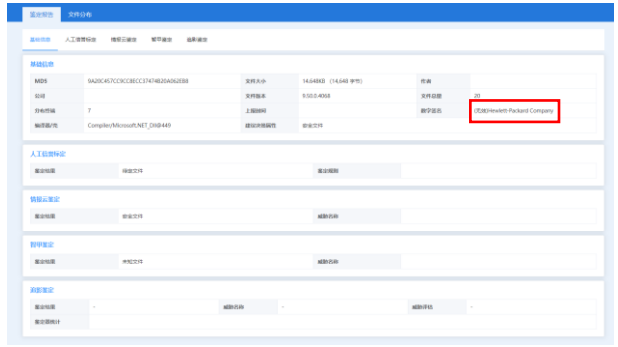
在线联网检测结果验证

离线场景

提取的证书元数据及微软判定签名结果



撤销库中证书元数据(序列号)



安天EDR产品签名结果显示

盗用证书签名逃逸检测的攻击识别

场景

- 证书被攻击者盗用，使恶意执行体规避了系统签名的验证机制

基于元数据支撑盗用证书的执行体识别

- 安天AVLSDK可对执行体进行元数据提取，得到身份相关的元数据信息，包括指纹，签发者信息等，结合恶意代码检测特征库，验证证书是否被盗用

微软签名检测结果及
签发者身份信息信息

2011年Winnti组织

公司	国家
ESTsoft Corp	韩国
Kog Co., Ltd.	韩国
LivePlex Corp	韩国/ 菲律宾
MGAME Corp	韩国
Rosso Index KK	日本
Sesisoft	韩国
Wemade	日本/韩国/美国
YNK Japan	日本
Guangzhou YuanLuo	中国
Fantasy Technology Corp	中国
Neowiz	韩国

Winnti组织利用/盗用证书

安天智甲分析检测结果

“低信誉”证书识别

场景

- 证书未被盗用，但证书签发的执行体存在风险行为，如合法证书签发的广告件

基于元数据支撑拥有合法签名的广告件识别

- 利用端点产品侧安天智甲防御检测系统，可以对执行体进行证书身份相关的元数据提取，包括指纹，签发者信息，厂商信息等，结合历史对该证书签发的执行体的判定信息，识别该证书处于低信誉值状态，结合恶意代码检测特征库，判定该执行体为广告件

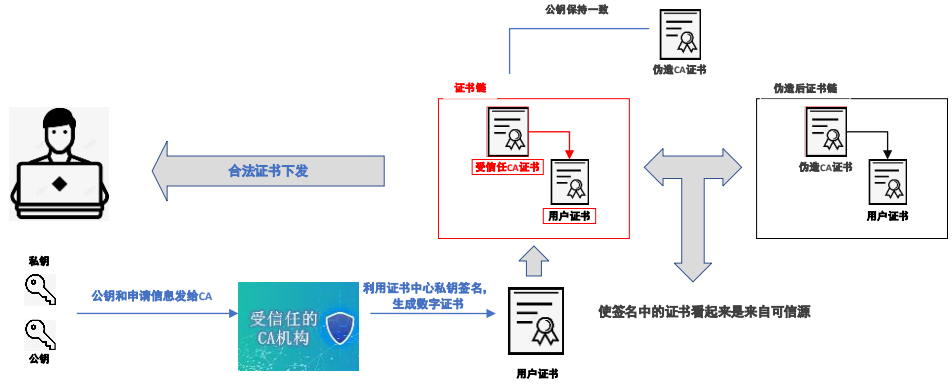
微软签名检测结果及
签发者身份信息数据

智甲分析检测结果

伪造证书的识别

场景

- 证书验签算法存在漏洞风险，攻击者可利用漏洞（如CVE-2020-0601）进行证书的伪造，攻击者利用证书签名算法（如椭圆算法），构造了欺骗性的代码签名根证书，对恶意代码进行签名，从而使该文件看起来合法



基于元数据识别伪造证书

- 安天引擎可以对执行体所携带的数字签名证书进行证书算法相关的元数据提取，包括签名算法，签名哈希算法等。可基于加密算法技术原理判断证书的签名算法是否与标准算法匹配，如椭圆ECC算法中所提供的G（椭圆曲线上的一个点）是否与标准证书算法匹配，从而判定是否为伪造证书

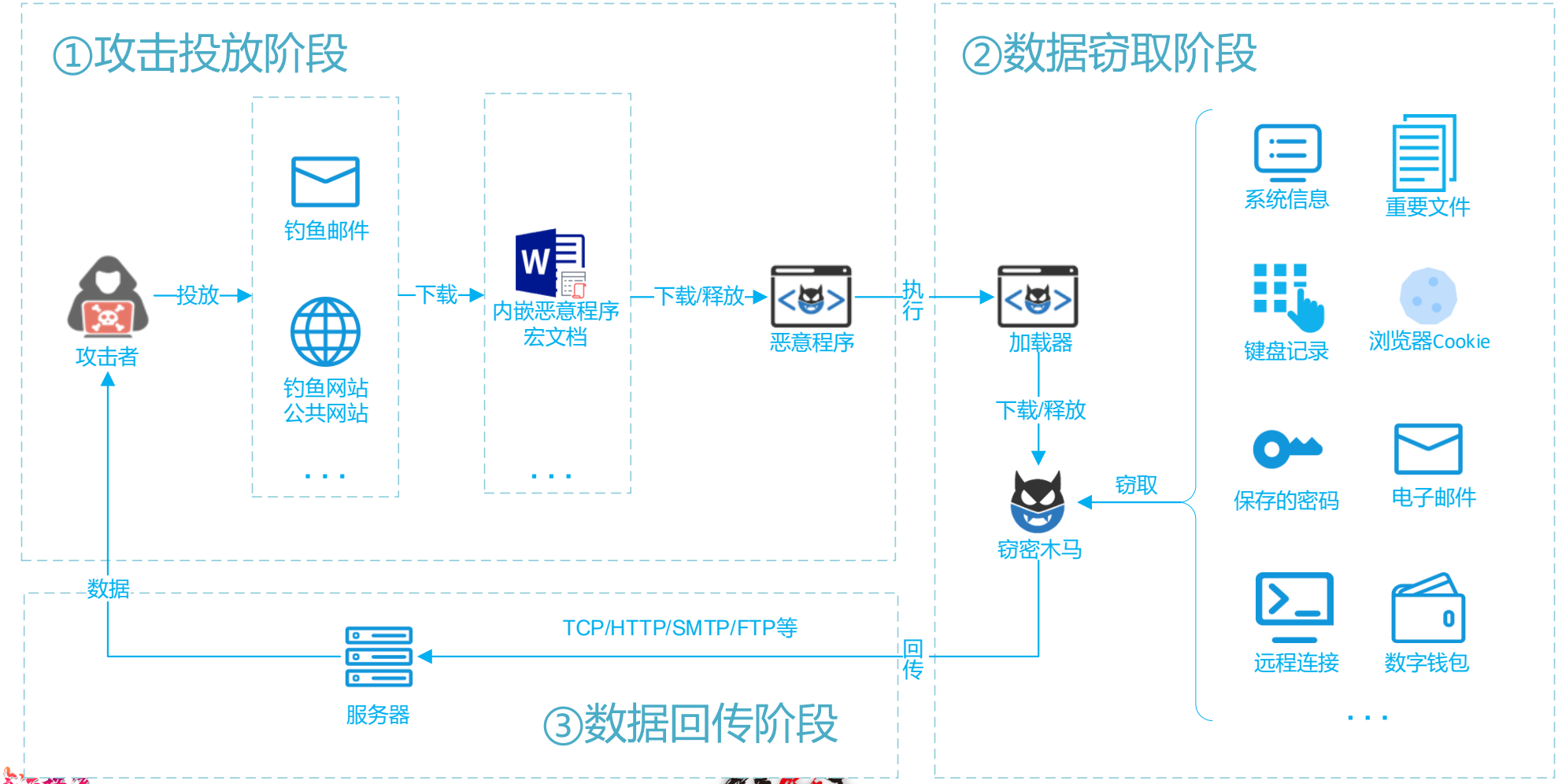
为什么要将复合文档元数据化

- 用户的**敏感和隐私性信息的载体**，不适合完整采集分析，将源文件发给厂商进行分析
- **支撑不提交文档内容情况下的脱敏分析**、标记和溯源等
- 客户防御场景的**主要风险入口之一**，存在多种利用复合文档的攻击方式

序号	威胁类型	威胁分类
1	利用宏攻击	公式宏
2		公式宏隐藏
3		XLSB公式宏编码
4		VBA宏
5	利用模版攻击	远程模版注入
6	利用内嵌对象攻击	内嵌DDE
7		内嵌OLE
8		内嵌钓鱼URL
9		内嵌ShellCode
10		多重嵌套
11	利用漏洞攻击	CVE漏洞
12	利用加密攻击	文档加密

基于复合文档的攻击手段

利用内嵌恶意宏的文档进行鱼叉邮件攻击



复合文档检测面临的挑战

攻击者利用复合文档的易篡改、易嵌套、易填充等特性躲避检测

1、通过元数据域掩藏方式躲避检测

在属性域内嵌恶意URL地址，利用正常VB代码读取属性域中的恶意URL，下载执行，躲避检测

2、通过多层嵌套方式躲避检

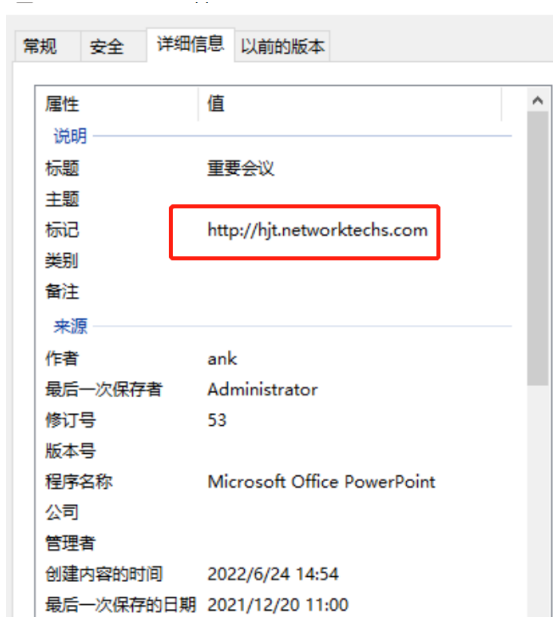
将恶意URL拆分成多个字串，将其嵌套在不同sheet表不同单元格中，利用正常的vba宏，读取拼接内嵌在单元格中的url，下载执行，躲避检测

3、通过多个cve漏洞躲避检测

同时包含3个相近的 cve, CVE-2018-0798 / CVE-2017-11882/ CVE-2018-0802等，通过不同cve触发漏洞执行，在某个cve被补丁修复难以触发，但可尝试触发其他漏洞

4、通过加默认密码躲避查杀

office有一个内置的默认密码，攻击者基于这点通过对恶意文档增加默认密码，来改变office的原始结构，导致杀软认为该文档加密而跳过检测



属性域中的标记被恶意填充恶意URL

```
00003568  00 00 00 00 00 00 00 00  00 00 00 00 00 00 00 00  ................
00003584  1C 00 00 00 02 00 A8 C3  44 17 00 00 00 00 00 00  ......¨ÃD.......
00003600  48 90 5D 00 6C 9C 5B 00  00 00 00 00 66 FE 01 DA  H.].l[.....fþ.Ú
00003616  BC 0A 01 11 22 33 44 F8  00 00 B9 34 6F 1D 8A B8  ¼..."3Dø..¹4o.Š¸
00003632  08 D2 58 8A 31 C1 8B 09  8B 49 14 83 C1 40 FF E1  .ÒXŠ1Á‹.‹I.ƒÁ@ÿá
00003648  37 65 30 37 39 61 32 35  32 34 66 61 36 33 61 35  7e079a2524fa63a5
00003664  35 66 62 63 66 65 9B 15  45 00 00 00 E8 FF FF FF  5fbcfe›.E...èÿÿÿ
00003680  FF C2 5E 83 C6 11 33 C9  66 B9 BB 16 80 36 B6 46  ÿÂ^ƒÆ.3Éf¹».€6¶F
00003696  E2 FA 5F C2 BE B6 B6 E3  3D 5A E7 E7 E5 E0 3D C3  âú_Â¾¶¶ã=Zççåà=Ã
```

使用了01的标记用于进入到cve-0798的流程，通过cve-0798进入对应的cve-11882

复合文档的细粒度元数据提取

序号	类别	字段	来源	描述
1	身份信息	文档作者	属性	文档编辑者，文档的归属
		归属公司	属性	文档所归属的公司
		内嵌邮箱	正文	正文中包含的邮箱号(来源、引用、目标对象等)
2	隐私信息	语言	属性	编写文档正文所使用的语言(中文、英文、俄文等)
		包含账号、卡号	正文	正文中包含的银行卡、登录账号等隐私账号信息
3	环境信息	依赖解释器名称	属性	打卡文档需要使用的软件名称(如Office,PDF等)
		依赖解释器版本	属性	打卡文档需要使用的软件版本(如Office2016)
		依赖运行平台	属性	执行文档的平台(如windows)
4	成分信息	编码方式	结构	文档编码的方式(如Office常采用XLSB编码)
		衍生数据类型	结构	包含的所有子文件个数，以及格式类型
		内嵌的OLE对象	结构	内嵌在文档中的其他对象数据(如内嵌在PDF中的JS)
		内嵌宏代码	结构	内嵌在文档中的宏代码
		内嵌的网络资源	正文	内嵌在文档正文中的URL、Domain、IP等
5	用途信息	文档主题	属性	文档属性中标记的主题
		文档标题	属性	文档属性中标记的标题
		敏感词	正文	文档正文中包含的敏感关键词列表
6	脆弱信息	CVE漏洞	结构	包含的CVE漏洞编号列表
		包含shellcode	数据	可疑的shellcode数据
		文档破损	结构	文档结构破损，无法正常打开
7	异常信息	文档加密	数据	文档通过密码加密，需要密码才能打开
		内嵌可执行文件	结构	文档内嵌了一个可执行文件(如包含exe文件)
		内嵌钓鱼、挂马URL	数据	文档内嵌了钓鱼URL
		属性域包含代码	属性	文档属性中包含代码
		正文包含脚本	正文	文档正文包含脚本
		宏自启动	行为	文档打开或关闭能自启动宏
		内嵌脚本加密混淆	行为	内嵌的脚本进行了加密混淆
		设置安全等级	行为	内嵌的脚本修改文档安全等级
		设置保护视图	行为	内嵌的脚本修改文档保护视图
		调用Active控件	行为	内嵌的脚本调用Active控件进行系统操作
8	信誉信息	信誉度	综合	文档的信誉度，可信、恶意、未知等
		威胁度	综合	文档的威胁度，低危、中危、高危等


```
{
 "Basic_info":{
  "file_path":"D:\文档\出行表.xlsx",                                  //文件路径
  "md5":"1B623586942D52CF16010232E14ED9B4",                          //文件MD5
  "size":189106,                                                     //文件大小
  "file_type":"Document/Microsoft.PPTX[:PowerPoint 2007-2012]",      //文件类型
  "password":true                                                    //是否加密
 },
 "identity_info":{
  "author":"Rabota",                                                 //作者
  "company":"",                                                      //公司
  "email":"minaq@email.com",                                         //邮箱
  "keywords":["重大","违纪","案件"]                                    //敏感关键字
 },
 "usage_info":{
  "theme":""案件纪录",                                               //主题
  "title":"统计信息"                                                 //标题
 },
 "composition_info":{
  "ole":[{                                                           //内嵌OLE信息
         "file_name":"mimikatz.exe",                                 //内嵌文件名称
         "md5":"4AE43708EB0DA8FB1F06B1D06BFA8F4F",                   //内嵌文件MD5
         "type":"BinExecute/Microsoft.PE[:X86]"                      //内嵌文件类型
         }],
  "url":[                                                            //内嵌URL信息
         "https://koucaloumantre.ml/ff",                             //钓鱼URL
         "https://ypsikaminos.gr/png.php"
        ],
  "macro_type":"vba"                                                 //宏类型
    }
}
```


某复合文档元数据示例

客户场景下的文档脱敏分析与追溯

安天智甲管理中心

- 终端资产管理
- 终端安全查杀
- 终端合规检查
- 可执行体治理
- 历史威胁溯源分析
- 远程取证分析
- 可视化运维管理
- 访问控制管理

智甲全网追溯

上传异常、成分、脆弱性元数据

上传身份、隐私、用途等元数据

安天智甲客户端集群

监控—新增

复合文档

服务器

.....

PC机

主动投递

敏感文档

03

安天协同用户达成全量执行体识别和精细管控

安天协同用户达成全量执行体识别和精细管控

信誉知识

执行体线索

用户

网络治理清晰化 | 客户优先的双重信誉机制 | 自维护的私有知识运营闭环

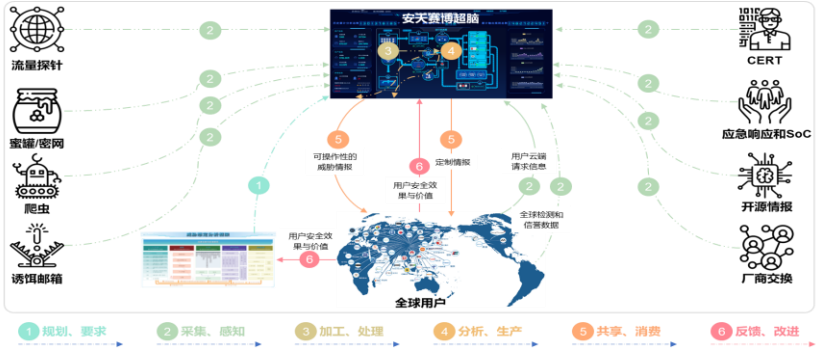
清晰化的网络治理与更有效威胁对抗能力

安全产品

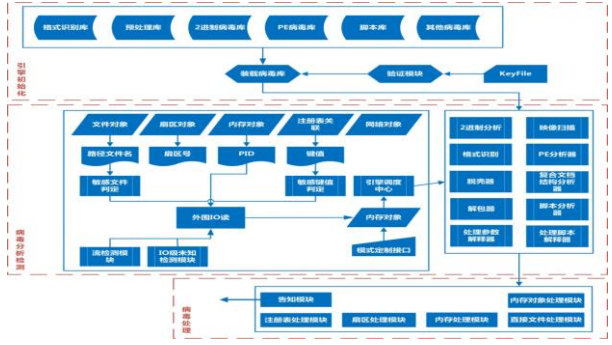
传统反病毒规则 | 向量级规则 | 知识标签

全量执行体元数据化能力

从特征工程向知识工程演进的后端工程体系



威胁对抗工程体系



安天AVLSDK威胁检测引擎

感谢大家的关注

安天冬训营 wtc.antiy.cn