**前言**

我们本次课程的内容主要讲解的是文档类恶意程序的快速判断技巧，涉及Word文档、PDF文件、HTML文件等等。判定的主要技巧很多时候还是依靠我们上次课程中的方法，也就是如果目标文件写得很不清晰很不规范，就说明这是恶意程序的作者故意让我们无法分析的，故意要躲避杀软的查杀的，那么我们就需要报毒了。而文档类恶意程序也会有它本身的一些特色，这个我会在这次的课程中依据实例给大家分析一下。

**四段命名法**

这里首先给大家复习一下我们的四段命名法。不同的杀软厂商会有各自不同的命名方法，但是我个人认为，只有四段命名法才是最为严谨，最能体现目标恶意程序本质的命名方法。四段命名法的命名规则如下：

恶意程序的类型 . 所运行的系统类型 . 恶意程序家族名称 . 变种号

其中，第一段是恶意程序的类型，比如如果是病毒，就是 Virus，如果是很宽泛意义上的木马，就是Trojan，如果是下载者木马，就是Trojan-Downloader，如果是漏洞利用程序，就是Exploit等等。第二段是所运行的系统类型，如果是 Windows的32位恶意程序，就是Win32，64位的就是Win64等等。第三段是恶意程序的家族名称，这个名字一般由病毒分析师来决定，可能恶意程序的作者也会在自己的病毒木马中给自己取个名字，但是这个名字一般不会被采用，或多或少都会在经过一定的修改后在使用。最后的变种号则是用于标识当前样本是这个家族的第几号样本，比如最先发现的就是a，之后的变种就是b，依此类推，一直到z。然后接下来就是aa、ab等等。比如对于这个恶意程序来说：

Trojan.Win32.Setiri.b

它的意思是，这是一个木马程序（Trojan），运行在 Windows的32位平台（Win32），家族名称是Setiri，是这个家族中的第二个样本（b）。再比如这个程序：

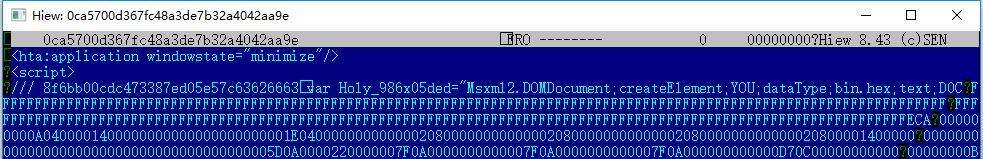
not-a-virus:Adware.Win32.Agent.z

这个程序的名称以not-a-virus作为开头，说明它并不是传统意义上的恶意程序，介于干净的程序和恶意程序之间的地带，这类程序主要分为广告程序、黑客工具等，而当前的这个程序则是Adware，即广告程序。这类型程序的命名方法是与恶意程序有点不同的，希望大家注意一下。

**样本分析**

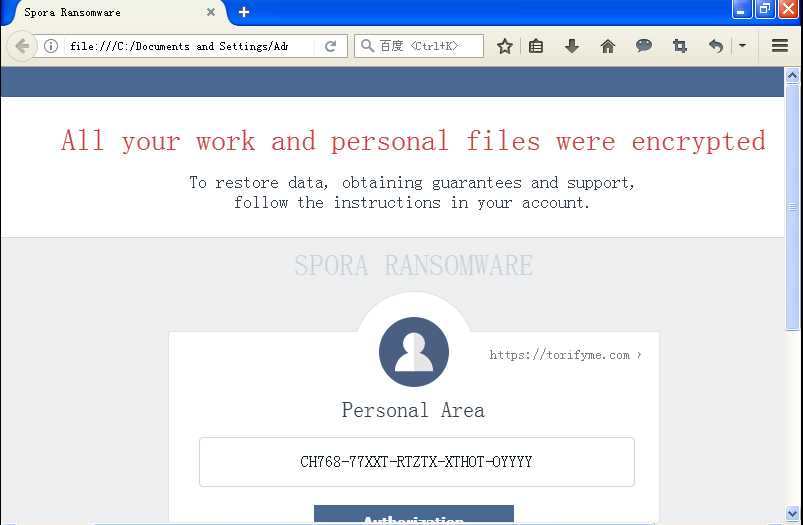
本次课程的文档类样本都是非常典型的，也是我们分析师几乎每天都会遇到的类型，每个样本都有它们各自的特点，是值得我们注意的。

**1、0ca5700d367fc48a3de7b32a4042aa9e**



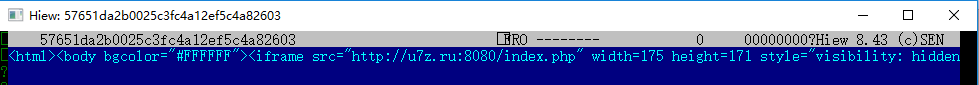
首先由最开始的HTA字符就可以知道，这个程序是一个HTA程序，再由后面的windowstate=”minimize”可以知道，它会以最小化窗口的形式运行，那么这个就是第一个可疑的地方，因为正常的这类小程序，没必要以这种隐藏的姿态执行。再由程序的主体部分，可以看到它很可能是采用了某种加密手段来阻碍我们的分析，依据这两点，就可以判定它是一个有问题的程序了。提取特征的话，可以直接在乱码的位置提取即可。

为了查看它到底是属于那种恶意程序，这里我们不妨在虚拟机里面运行一下，运行前需要把它的后缀，也就是.hta加上。运行后可以看到，它首先会弹出一个写字板，由这里显示出来的文字可以知道，这个恶意程序很可能是针对于俄罗斯进行攻击的。不久后就弹出了这样一个网页：



内容是说你的文件都被加密了，并且由这个网页的标题也可以知道，这是一个类型为Spora的勒索程序，由于这个勒索程序家族很出名也很庞大，因此我们不妨就直接采用这个名字作为家族的名称，可将这个恶意程序归类为：Trojan-Ransom.HTA.Spora.a。

**2、57651da2b0025c3fc4a12ef5c4a82603**

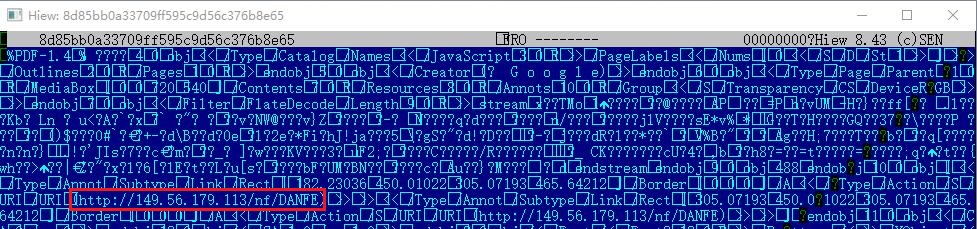


首先由最开始的html就可以知道它是一个网页文件，然后由iframe以及后面的visibility:hidden，说明这个程序会以隐藏的方式访问一个网站，这就是非常令人怀疑的地方了，通过这些，基本就可以确定它是恶意的了。当然了，严谨起见，还应该好好分析一下这里出现的这个网址究竟是不是恶意的，如果是的话，还需要把这个网址拉黑。但是一般来说，直接把这个程序判定为是恶意的并不会有什么问题。

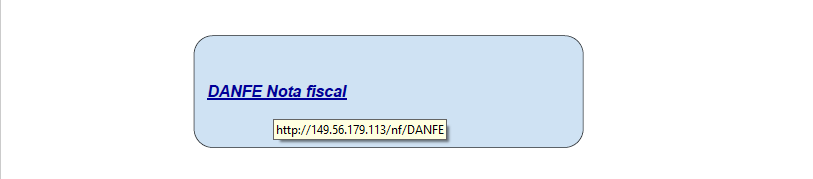
一般来说，这种iframe类的程序，要么是采用visibility:hidden的方式达到隐藏访问的目的，要么是采用将窗口大小都设置为0的方式实现，即将width和height全都设置为0，那么也能够达到相同的目的。对于这类的恶意程序，如果采用手工一个一个去拉黑的方式进行对抗，那么很显然是不现实的，因此应该引入高级查杀的方式来应对这一类的恶意程序，比如静态启发。关于静态启发，我会在高级篇的课程中跟大家详细讲解，这里可以跟大家讲一下大概的思路。对于本程序而言，如果发现程序中出现了iframe，并且后面还跟有visibility:hidden或者width=0，height=0这样的字符，那么就可以判定为这是可疑程序了。

最后，可以将这个程序归类为：Trojan.HTML.iframe.a。

**3、8d85bb0a33709ff595c9d56c376b8e65**

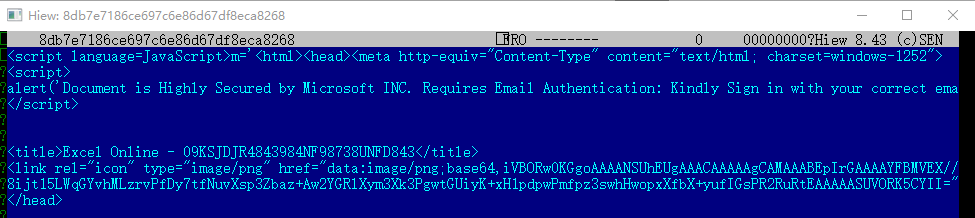


首先由程序最开始的位置可以知道，它是一个PDF文件。对于这类的文件，一般来说是难以直接通过阅读它的代码来得知它的问题的，不过这里可以看到一个网址，姑且列为怀疑的对象。然后不妨将这个文件的后缀修改为.pdf，运行看一下：

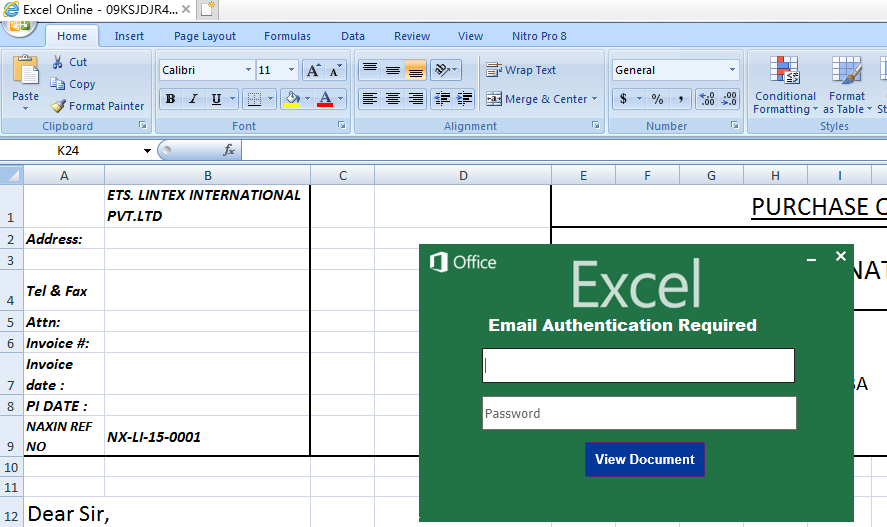


可以看到，这个PDF文件内容只有一个网页的链接，这个网址和我们刚才在文件中的网址是一样的。那么这就是可疑之处，因为你作为一个PDF文件，按道理会包含有一定的信息，可是你现在仅仅作为一个跳板来诱导用户去访问另一个网站，这就很不正常了。一般来说，这种网站以钓鱼类的居多。当然了，严谨起见是要仔细分析一下这个网站的，但其实分析至此，我们就基本可以判定这个PDF文件是恶意的了。可以将其归类为：Trojan.PDF.Phish.a。

**4、8db7e7186ce697c6e86d67df8eca8268**

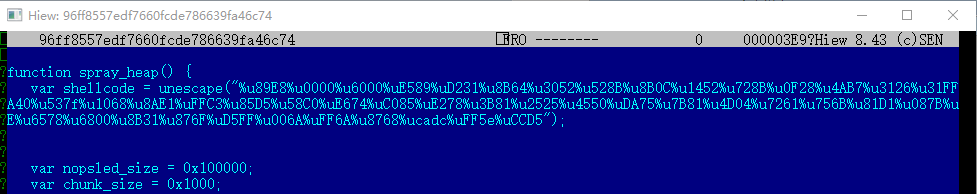


首先可以判断出它是一个HTML的文件，并且文件内容中也是包含有大量的加密内容。由alert后面的内容可知，它会要求用户在输入电子邮件的账号和密码之后，再获取相关的文档，这就是非常值得怀疑的地方了。要求我们输入账号还算正常，但是连密码也要输入进来，那就很奇怪了，它有钓鱼程序的嫌疑。不妨将其后缀加上.html，然后运行看一下：



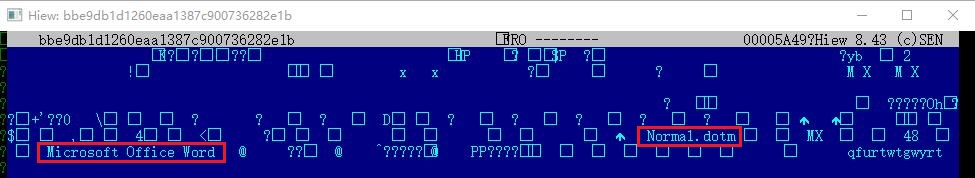
可以看到，它以一个表格图片作为背景，诱惑用户输入账号密码来获取完整的表格文件，分析至此，就可以将其归类为：Trojan.HTML.Phish.a。

**5、96ff8557edf7660fcde786639fa46c74**

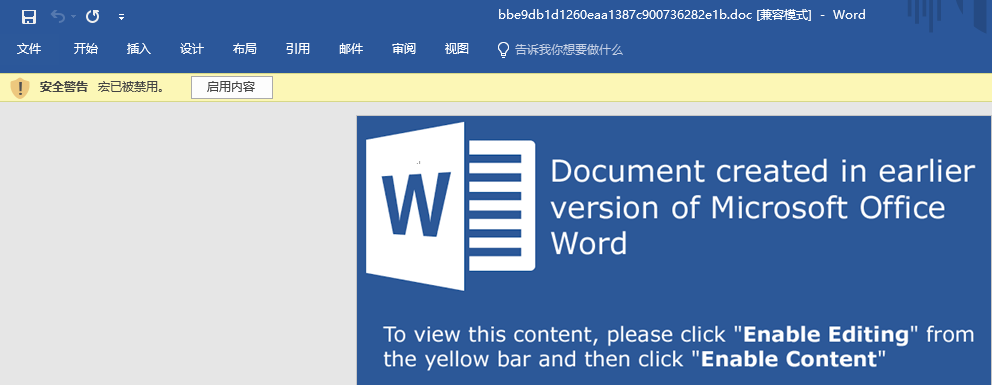


首先这里面出现了PDF，可以知道它是一个PDF文件或者是PDF文件的一部分，还可以知道这里面包含有JS的脚本文件。从这里的spray\_heap字符，我们有理由怀疑，这个文件使用了堆喷射技术，而且这类编码在文件中还有很多，那么基本就可以认为这是一个利用漏洞进行攻击的PDF文件了。于是我们就可以将其归类为：Exploit.PDF.Agent.a，当然了，如果将其定义为JS类型也是可以的，毕竟漏洞利用的代码是以JS脚本的形式编写的。

**6、bbe9db1d1260eaa1387c900736282e1b**

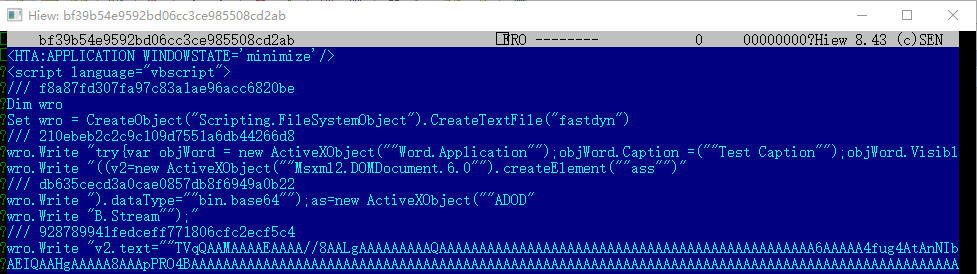


这个文件在文件头部的位置并没有明显的用于表明文件类型的字符，但是如果我们往下翻页查找，可以看到Microsoft Office Word以及Normal.dotm这样的字符，那么就可以知道它是一个Word文件了。一般来说，Word文档类的恶意程序以内置宏代码的下载者文件居多，各个杀软厂家都会有自己的解密工具来将宏代码提取出来，第三方也有一些比较好用的工具来实现这样的功能。或者我们直接利用hiew也可以找到宏代码，但是比较麻烦，也不直观。但是我们这里并不需要对样本进行详细分析，而是快速判断，所以不妨将文件的后缀，也就是.doc加上，尝试运行一下：



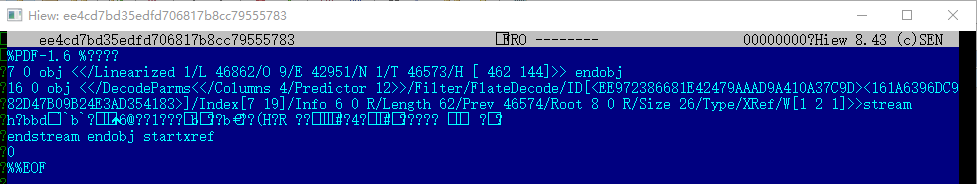
首先可以看到，我们的Office软件已经将宏禁用了，说明这个文件中确实包含有宏。再从文档的内容可以看到，这个恶意程序的作者为了诱惑用户运行这个宏，会骗用户自己去手动解除宏或者编辑功能的禁用状态，从而实现恶意的目的。那么由此就可以知道，这是一个恶意文件了。可以将其命名为：Trojan.MSWord.Agent.a。

**7、bf39b54e9592bd06cc3ce985508cd2ab**

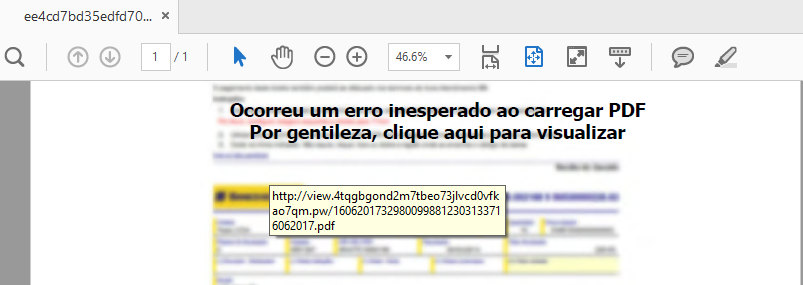


首先由文件开头可以知道它是一个HTA文件，再由后面的windowstate=”minimize”可以知道，它会以最小化窗口的形式运行，并且在文件主体中，我们也发现了大量的采用BASE64加密的内容，这对于代码的阅读就显得很不直观。所以不妨再利用虚拟机运行一下，可以发现它其实就是我们最开始分析的那个样本的变种，因此可以将其命名为：Trojan-Ransom.HTA.Spora.b。

**8、ee4cd7bd35edfd706817b8cc79555783**



首先可以看出来这是一个PDF文件，但是从代码中似乎并没有什么问题，所以这里不妨将其后缀修改为.pdf然后再运行看看：



可以发现，这个文件运行的背景也是一个很模糊的表格，当我们把鼠标移动到这个表格上的时候，就会出现一个网址，也就是当我们点击鼠标的时候，就会直接访问这个网址。这同样是一个非常可疑的地方，很多时候，恶意程序的作者就是利用一些公司的后勤或者财务人员的薄弱的安全意识进行攻击的，诱惑这些人员只有点击了这个图片才可以获取到这个文档。所以我们可以将这个网址加入到杀毒软件的数据库，从而实现拉黑。可以将这个文件命名为：Trojan.PDF.Phish.b。

**小结**

我们这次介绍了常见的文档类恶意程序的快速判断的方法。这些分析技巧虽然简单，但却非常有效。事实上，很多文档类恶意程序就是利用大家的安全防范意识薄弱，才有的可乘之机，所以安全的道路，依旧是任重而道远。