doi 10.3969/j.issn.1671-1122.2015.09.041

基于社会工程学的邮件样本关联分析

梁宏[1]，张慧云[2]，肖新光[2]
（1. 国家计算机病毒应急处理中心，天津 300457 ；2. 安天科技股份有限公司，黑龙江哈尔滨 150028）

摘 要 ：文章从基于社会工程学的邮件攻击方式和造成的危害出发，对利用社会工程学的电子邮件攻击进行了深入的分析。目前利用社会工程学技巧依托电子邮件发起攻击已是常见攻击方法，是当前恶意代码流行的重要手段。文章依托捕获的一些安全事件，对邮件的传播手段、附件文件的攻击方法进行了关联分析。最后，通过提取同类特征挖掘出更多的类似攻击邮件，并进行了整体的关联分析与总结。

关键词 ：社会工程学 ；电子邮件 ；恶意代码 ；样本
中图分类号 ：TP309 文献标识码 ：A 文章编号 ：1671-1122（2015）09-0180-06

中文引用格式 ：梁宏，张慧云，肖新光 . 基于社会工程学的邮件样本关联分析 [J]. 信息网络安全，2015，(9) ：180-185.

英文引用格式 ：LIANG H, ZHANG H Y, XIAO X G. Analysis of E-mail Sample Correlation Based on Social Engineering[J]. Netinfo Security, 2015, (9) : 180-185.

Analysis of E-mail Sample Correlation Based on Social Engineering

LIANG Hong[1], ZHANG Hui-yun[2], XIAO Xin-guang[2]
(1. *National Computer Virus Emergency Response Center, Tianjin* 300457, *China*; 2. *Antiy laboratory, Harbin Heilongjiang* 150028, *China*)

Abstract: Starting from the email attacks and hazards of view based on social engineering, the paper deeply analysis email attacks based on social engineering. Currently, email attack based on social engineering techniques is a common method of attack, and is an important channel of the malicious code. This paper relies on a number of captured events, and correlative analysis attack method by means of communication and attachment file of email. Finally, by extracting similar characteristic, the paper digs out more similar attack email, and gives the overall correlation analysis and summary.

Key words: social engineering; E-mail; malicious code; sample

0 引言

2014 年，对互联网影响重大的安全案例层出不穷，“ 心脏出血 ”（Heartbleed）漏洞影响了数以万计的服务器，敲诈者病毒、伪银行木马让上百万用户陷入困境，社交网络钓鱼真假难辨，垃圾邮件的数量持续攀升，网络安全状况日益严重。调查显示，2014 年 30.4% 的用户遭遇过网络钓鱼，这也是传统 PC 和移动终端共同面临的安全问题。随着智能手机等移动终端的普及，网络钓鱼已经逐步向移动终端转移，形成了“ 网络 + 社交 + 电话 ” 的复合模式。病毒和木马的传播方式也更多地利用社会工程学的方法，通过结合用户的使用习惯、身份信息等构造出仿真度极高的网站、电子邮件、即时通讯消息等诱骗用户。越来越多的安全事件有了社会工程学的因素，用户端往往是整个安全体系中最薄弱的环节，即便再强大的技术也难

收稿日期 ：2015-07-15
作者简介 ：梁宏（1977-），女，天津，高级工程师，硕士，主要研究方向 ：计算机病毒应急处置 ；张慧芸（1988-），女，黑龙江，工程师，主要研究方向 ：恶意代码分析 ；肖新光（1974-），男，黑龙江，高级工程师，主要研究方向 ：反病毒引擎、恶意代码分析、APT 检测与分析。
通讯作者 ：梁宏 liangh@cverc.org.cn

以阻止安全事件的发生[1-3]。也正因为如此，社会工程学在安全事件中大行其道。

社会工程学攻击在安全圈被称为社交工程，是一种通过对受害者本能反应、好奇心、信任、贪婪等心理陷阱采取如欺骗、伤害等危害性手段，获取利益的手法。传统的攻击者在系统入侵的环境下存在很多的局限性，而社会工程学攻击则通过利用人为的漏洞缺陷进行欺骗来获取系统优势，进而获取系统控制权。这种攻击表面上难以察觉，不需要与受害者目标进行面对面的交流，不在系统留下任何可被追查的日志记录，难于追查。

电子邮件是个人工作生活的重要应用工具，绝大多数政府、企业的办公仍依赖于邮件交互。电子邮件系统一旦被攻击者攻破，轻则泄露个人资料，重则泄露国家机密，损失不可估量。近年来，电子邮件也成为黑客最易取得成效的 APT 攻击入口。在某些案例中，攻击者会利用受害者的电子邮件账号来增加他们鱼叉式网络钓鱼攻击邮件的可信度。而经过足够多的研究分析后，网络犯罪分子可以制造出社交工程诱饵，骗取足够多的用户点击网络链接或者打开邮件附件。2011 年的 RSA SecurID 攻击事件就是以电子邮件为切入口，最终入侵了 RSA 开发用服务器。因此防范针对电子邮件发起的攻击就非常必要[4-6]。

目前，利用社会工程学技巧依托电子邮件攻击已是常见攻击的方法，已有相关的研究对其进行分析。采用社工技巧大规模传播恶意代码的攻击方式成为恶意代码流行的重要通道[7-9]，本文依托近期捕获到的一些事件对此进行关联分析。

首先通过诱饵信箱持续捕获大量利用社工技巧进行传播的带毒邮件，随后展开对某类邮件攻击手法和技巧的批量分析。从诱饵信箱中随机抽取两封附件攻击手段一致的邮件作为分析起点。第一封邮件的捕获时间为 2015 年 4 月 11 号，这是一封伪装成摩根大通集团的钓鱼邮件，邮件包括一个 zip 压缩包和一个空文档文件（空文档文件是由于发送与接收的客户端不一致引起的邮件客户端解析乱码行为），压缩包解压后是一个 PDF 图标的 PE 文件，运行文件后会从网络下载其他文件 ；另一封邮件的捕获时间为 2015 年 4 月 13 号，这两封邮件的附件行为一致。首先对第一封邮件进行分析，随后对两封邮件的传播手段、附件文件的攻击方法进行关联分析，最后通过提取类似特征，挖掘出更多的类似攻击邮件并进行整体的关联分析与总结。

1 邮件 1 分析

原始邮件 1 内容为英文，主要内容是提示收件人了解在线账户的最新情况，并声明邮件包括附件内容都是保密的，仅发送给个人使用者。

1.1 邮件1内容分析

表1 邮件元数据提取

邮件主题	JP Morgan Access Secure Message
发送时间	2015/4/11 (周六) 2:24
发件人	JP Morgan Access [service@jpmorgan.com]
附件 1 名称	JP Morgan Access – Secure.zip
附件 2 名称	未命名的附件 00273.txt

如表 1 所示，该邮件假冒摩根大通集团发送，利用社会工程学进行攻击。邮件恶意代码的社会工程学的攻击手段主要是通过发送带有欺骗、伪装等主题内容的邮件，诱使受害者点击附件。该邮件利用摩根银行拥有庞大用户群这一事实，通过在线银行的账户文件的敏感话题为诱饵，引诱用户点击附件查看账单，并且特别提醒声称邮件内容为机密信息。如果用户并没有摩根的在线银行或者错误地接收到此邮件，看到机密信息内容出于好奇心也可能对此附件感兴趣。

邮件附件是 1 个下载者病毒，为 1 个 zip 压缩包，压缩包解压后是 1 个以 PDF 为图标且扩展名为 src 的文件，实际上是一个 PE 可执行程序，多数人对此并不了解，双击后，会自动运行 PE 文件。图 1 为邮件结构和行为链图。

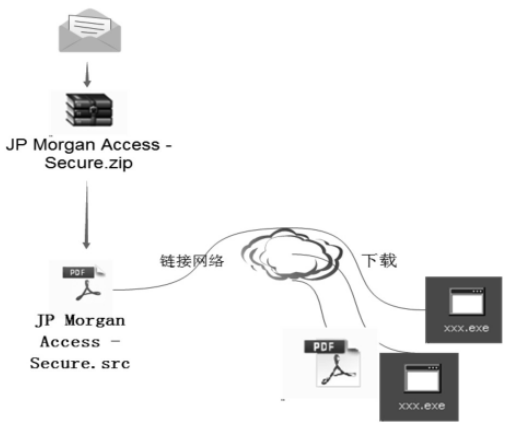

图1 结构和行为链

表2 邮件标签分析

病毒名称	Trojan[Downloader]/Win32.Upatre
原始文件名	JP Morgan Access – Secure.src
MD5	32E1F5DED6E9C573293BB6343F785A9F
处理器架构	X86–32
文件大小	24,064 字节
文件格式	BinExecute/Microsoft.EXE[:X86]
时间戳	2004–09–05 12:06:14 ！
数字签名	NO
加壳类型	无
编译语言	Microsoft Visual C++
VT 首次上传时间	2015–03–02
VT 检测结果	48 / 57
样本定性	Downloader

1.2 本地行为描述

1）概括描述，邮件 1 是长度仅有 24KB 的可执行文件，并具有反调试功能，经过一系列的解密将恶意代码解密执行。样本利用 Windows 消息机制创建隐藏窗口，将恶意代码写在自定义的函数中[10,11]。运行后，查看临时目录是否有进程副本，如果有继续运行，如果没有将自身复制到临时目录下并更名为 raturas.exe，邮件 1 运行过程中采取自修改指令方式，将关键代码以文件形式加密保存在自身进程中，运行后将加密数据进行解密后开始重新加载自身进程后从网络上下载其他进程文件。

2）细节分析，恶意代码首先创建一个线程，通过消息机制进行线程切换，将线程放在消息自定义函数中执行。在线程中恶意代码首先将加密数据放在内存 810000 中。随后经过解密 1 代码段进行初步解密，如图 2 所示。将解密后的数据放在 820000 内存段，此时恶意代码已经完成了初步解密，在经过第二次解密通过 403CCC 完成，完成后的代码放在 83000 中，生成一个新的 PE 可执行文件。

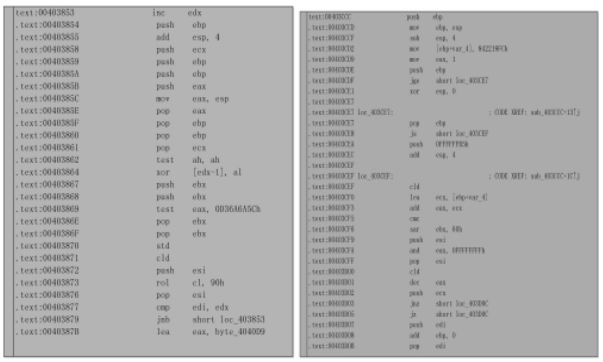

图2 左侧为解密1代码段，右侧为解密2代码段

当完成解密后，该恶意代码进行创建子进程，在子进程中执行申请空间、写入代码、执行代码等功能。随后子进程开始运行并连接网络。

1.3 网络行为描述

恶意代码运行后与远程服务器通信，进行文件下载。并访问如表 3 所示。

表3 恶意代码远程通信

域名	IP	GET	result
checkip.dyndns.org	**********		本机 IP
milbrookelt d.co.uk	192.185.86.160	/cufon/sdocn.pdf	失效
nationalpalletdelivery.com	192.185.86.183	/demo/documentation/sdocn.pdf	失效
	190.111.9.129		失效

恶意代码运行后首先在系统临时目录下创建文件名为 r657temp.log 的 TXT 文件，内容为 ：

C :JP Morgan Access-Secure.exe

随后进行调用删除命令进行文件删除，该 PE 文件是样本创建的副本。

恶意代码链接 checkip.dyndns.org 域名，该域名有获取本地 IP 地址的功能，恶意代码是利用该域名确定此时网络连接是否正常。获取本地 IP 后，进行 milbrookelt d.co.uk、nationalpalletdelivery.com 域名的链接进行下载指定文件。最后进行 IP（190.111.9.129）地址链接。循环进行此步骤。下图 3 为整个恶意代码的工作流程。

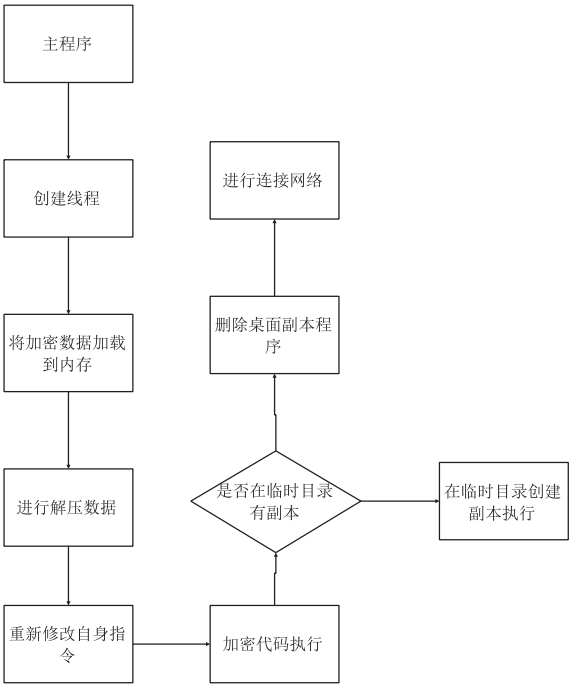

图3 程序流程图

1.4 行为分析

邮件 1 首先是以社会工程学攻击为前导，当用户点击附件后，将文件副本隐匿在临时目录下并进行连接网络操

作，样本是一个“ 下载者 ” 病毒，不涉及任何启动项功能。整体附件样本的代码技术简练而且有效。

2 邮件 2 分析

2.1 邮件2内容分析

表4 邮件元数据提取

邮件主题	Invoice*****
发送时间	2015/4/1412:24
发件人	******
附件 1 名称	Invoice_***.zip
附件 2 名称	未命名的附件 00625.txt

表5 邮件标签分析

病毒名称	Trojan[Downloader]/Win32.Upatre
原始文件名	Invoice_004AP71
MD5	6093329DBDA17782BB8DC31CF223A188
处理器架构	X86–32
文件大小	31232 字节
文件格式	BinExecute/Microsoft.EXE[:X86]
时间戳	2006–05–11 20:46:41
数字签名	NO
加壳类型	无
编译语言	Microsoft Visual C++
VT 首次上传时间	2015–04–13
VT 检测结果	41 / 56
样本定性	Downloader

2.2 本地行为描述

邮件 2 本地行为和邮件 1 相同。

2.3 网络行为描述

该附件运行后，连接网络，尝试下载两个文件到本机执行，其中，一个文件下载失效，另一个样本文件下载的是 PDF 文件，该 PDF 文件从后端下载并且打开后，让用户觉得附件是个真正的 PDF，然而当该 PDF 打开之时，附件文件已经完成了网络下载恶意代码功能。

该 PDF 文件的内容可以看出和社工邮件正文内容并不一致，这种情况可能是由于邮件攻击作者批量存放了多种攻击手法与多种主题的 PDF 文件，并在攻击时刻将邮件攻击的正文内容和对应的 PDF 文件下载地址关联混乱导致，但也不排除是作者手法粗略并没有考虑一致性所导致。

3 亲缘关系分析

两个邮件均是 2015 年 4 月捕获到的，通过对两个邮件的整体结构，以及附件代码的手法进行了亲缘关系分析，发现两个邮件的亲缘关系紧密，猜测属于同一作者进行的批量攻击，通过结构链、代码行为、代码结构三个方面的亲缘关系关联与对比。

通过比对表 1、表 2、和表 4、表 5、可以发现两封邮件的结构链几乎完全一致，均是一个压缩包，压缩包解压后为一个 PDF 图标的 PE 文件，运行后在网络上下载其他文件。

3.1 代码行为

如图 4 所示附件样本运行后均是在临时目录下创建副本。邮件 1 以 raturas.exe 为文件名，邮件 2 以 docfree.exe 为文件名，并且两个邮件均在临时目录下创建文本文件调用 CMD 命令行方式删除 C 盘副本文件，样本自删除的方式也相同。

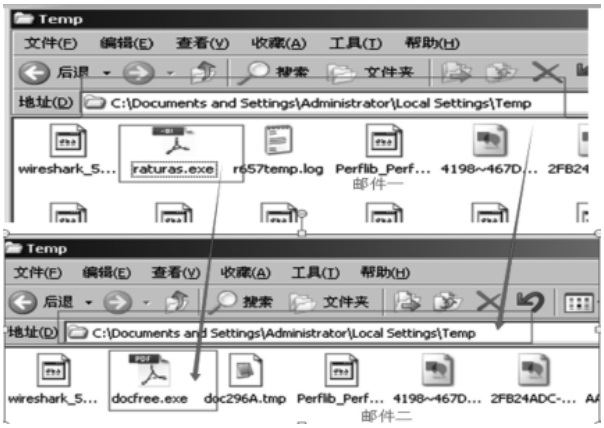

图4 样本衍生文件目录、图标相似

3.2 代码结构

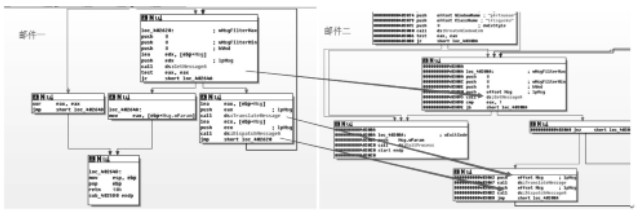

图5 消息机制

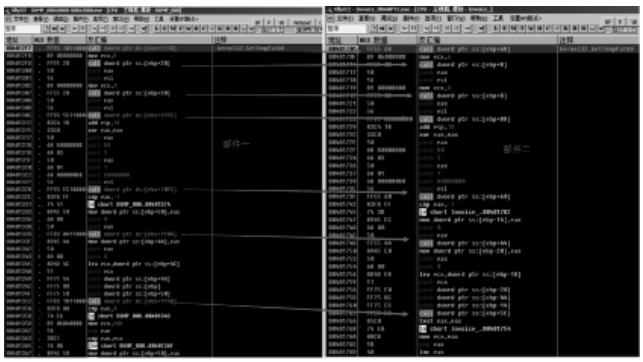

图6 解密后代码结构

两个样本均是以消息机制为整体代码编写方式，将恶意代码写在自定义函数之中，并且两个样本均是将代码解密出来后运行，并且所有 API 均是解密动态获取。如图 5 和图 6 所示两个样本的消息机制和代码结构几乎一致。

3.3 网络行为

邮件一

域名	IP	GET	result
checkip.dyndns.org	216.146.38.70		本机 IP
milbrookelt d.co.uk	192.185.86.160	/cufon/sdocn.pdf	失效
nationalpalletdelivery.com	192.185.86.183	/demo/documentation/sdocn.pdf	失效

邮件二

域名	IP	GET	result
checkip-iad.dyndns.com	216.146.38.70		本地 IP
kapil.amsinformatics.com	216.245.213.210	GET /images/monuk14.png	失效
syc.mxserver.ro	176.223.122.103	GET/wp-includes/images/monuk14.png	PDF 文件

图7 网络行为相似性

如图 7 所示两个样本均是首先访问 checkip-iad.dyndns.com 域名来测试网络连通状态。可见联网方式的亲缘关系也是一致的。

3.4 分析与结论

对两封邮件通过分析邮件结构、结构链、样本本地行为、网络行为、样本代码等多方面内容得知两封邮件具有高度亲缘性。

通过上述两封邮件的分析与关联，分析人员通过邮件的结构链，以及附件行为和代码结构等特征在库中抽取了更多的邮件进行批量分析。

3.5 社工手段

通过附件特征与结构的亲缘密切关系分析人员抽取了 110 个类似邮件，邮件附件的大小均在 23~32kb 中间，邮件的攻击手段有的是通过银行支票作为诱饵，有的是通过扫描件作为诱饵，经过统计对 110 个邮件进行攻击技巧标签化，共划分为 14 个标签 ：

1) 附件带 document 字样
2) 信用检测
3) 无正文内容
4) 固定签名
5) 问候
6) 公司薪资
7) 年度报告
8) 扫描件
9) 银行账单
10) 交易类
11) 汽车保险
12) 合作伙伴
13) 更新网上银行
14) 银行密码重置

重点说明 ：

1) 附件带 document 字样

该类邮件使用攻击的附件名称均包括 document 名称。而邮件的主题内容不一致，有的与银行、支票有关，有的是空内容，或者简短的内容但附件攻击代码一致。

2) 固定签名

图8 固定签名

通过分析发现，同一类邮件的攻击时间和攻击人基本是一致的，如图 8 所示，那么，我们的推断是否完全正确呢，下一节我们对 110 封邮件进行发送时间聚合分析。

如图 9 所示，附件为 document 的攻击邮件最多，其中是信用检测、无内容群发、固定签名、问候等也是常用的攻击手段。

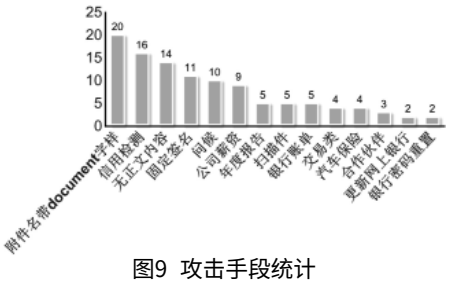

图9 攻击手段统计

3.6 攻击时间分析

通过对 110 封邮件的社工技巧手段与发送时间的分析发现，几乎所有采用同一社工技巧的邮件都是同一时间发送，如图 10 所示。

通过对 110 个邮件的时间摘取分析发现，同一类别邮件攻击者一般采取批量攻击，也就是说在相同时间点发送批量邮件进行攻击，而且通过时间发现，提取邮件的最早攻击时间是 2014 年，2015 年 3 月又开始进行批量攻击，到 2015 年 4 月再次进行大规模性攻击。

4 结束语

安全威胁如今已经成为 IT 业界的一种常态，每天都

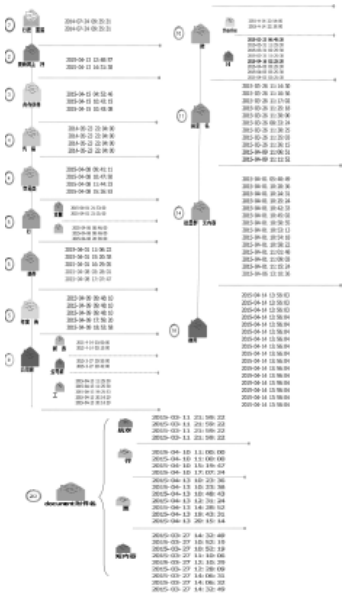

图10 提取所有抽取邮件的攻击时间

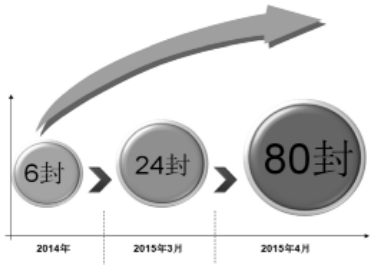

图11 攻击时间与数目

有将近数十万种病毒产生，数以万计的网络攻击发生，随着安全技术与产品的发展，以及安全产品的多级部署，普通终端用户的个人电脑安全状况已经有了普遍提升。因而，除了技术层面的角逐，恶意攻击者越发注重于结合社会工程学方法，打通人这一薄弱环节，从而绕过层层技术防控手段。社会工程学的融入加大了安全事件的复杂度，也增加了安全威胁的防范难度。

一直以来，电子邮件以其使用频率高、传播成本低等特点，成为恶意用户传播病毒、发动攻击的重要手段和有效通道。早在 1999 年爆发的 Happy99 病毒、梅丽莎（Melissa）病毒等作为邮件病毒的鼻祖，之后对照梅丽莎的“蓝本”又爆发了爱虫（LoveLetter）病毒、马吉斯（Magistr）病毒。这些病毒均属于“蠕虫”类型，具备自我扩散的能力，传播范围广，传播次数快，短时间快速传播，一方面会对网络带宽带来很大压力，也很容易被感知和发现。今天的邮件入口，已经完全告别了当时的“天真病毒时代”，邮件病毒从宏病毒和其他使用邮件 API 自动发送的二进制样本，到使用格式文档溢出的 APT 攻击等。攻击手法从普通的邮件攻击、钓鱼攻击、到鱼叉式钓鱼攻击。攻击手法从粗糙到精密，攻击目标从“扩散”到“扫射”再到“狙击”。

尽管此次分析的相关邮件并不具有高度定向性，但其呈现出小批量选择性发送的特点，但对受到欺骗的用户来说可能会造成一定的损失。在邮件中采用了较为精细的社工技巧后，必然有一定比例的用户会被欺骗，这种攻击对受害用户来说，达成了与定向攻击类似的效果。当浏览器、主机、服务器系统的安全在进一步的加固情况下，邮件对攻击者来说是个可以直达目标的上佳入口，这意味着通过邮件进行攻击的方法无论在高级威胁，还是类似扩散僵尸网络、敲诈其他的一般性的网络犯罪中，都将挥之不去。

网络安全是国家安全的重要组成部分。用户是网络活动中的主要参与者，也是最为关键的一环，互联网应用日新月异的同时，网络环境也日趋复杂，网络安全威胁的防范早已不能单单依靠技术，用户安全意识的提升对网络安全整体状况的改善至关重要，互联网的安全需要安全从业者与广大用户的共同努力。（责编 程斌）

参考文献：
[1] 曹麒麟 , 张千里编著 . 垃圾邮件与反垃圾邮件技术 [M]. 人民邮电出版社 , 2003.
[2] 赵晓丹，徐燕 . 垃圾邮件分类技术对比研究 [J]. 信息网络安全，2014，(2)：75-80.
[3] 百度百科 . 社会工程学 [EB/OL]. http://baike.baidu.com/link?url=qR8PYX92aCGVxUiNuEEgVzVCpVoIU2BNwWwUXEdw1T8bvYYqnE-oZP0zeFpoAiJ77GBqk0IjoX3y_aWwFMeF2K/20150610,2015-06-10.
[4] 杜雷，辛阳 . 基于规则库和网络爬虫的漏洞检测技术研究与实现 [J]. 信息网络安全，2014，(10)：38-43.
[5] 阿里云资讯 .RSA SecurID 受到攻击值得关注但可能不是一个致命缺陷 [EB/OL].http://www.aliyun.com/zixun/content/2_6_538635.html/20150610,2015-06-10.
[6] 向旭宇 . 邮件安全审计和过滤技术研究与实现 [D]. 长沙：中国人民解放军国防科学技术大学，2003.
[7] 宁戈，张涛，文伟平，等 . 一种非堆喷射的 IE 浏览器漏洞利用技术研究 [J]. 信息网络安全，2014，(6)：39-42.
[8] 第十四次全国信息网络安全状况暨计算机和移动终端病毒疫情调查分析报告 [EB/OL].http://www.cverc.org.cn/head/diaocha2014/report2015.pdf/20150610,2015-06-10.
[9] 关通，任馥荔，文伟平，等 . 基于 Windows 的软件安全典型漏洞利用策略探索与实践 [J]. 信息网络安全，2014，(11)：59-65.
[10] 周威成 . 信息过滤方法的研究与应用 [D]. 北京：华北电力大学（北京），2003.
[11] 叶嘉羲，张权，王剑 . 基于权限控制和脚本检测的 Webview 漏洞防护方案研究 [J]. 信息网络安全，2015，(3)：38-43.