网络钓鱼防御方法研究

倪天华,朱程荣
(同济大学 计算机科学与技术系,上海 201804)

摘 要:网络钓鱼主要是指利用互联网进行的一种欺诈行为。随着互联网的广泛普及,针对在线身份窃取的网络钓鱼活动日益加剧。阐述了网络钓鱼的基本概念;对现有的网络钓鱼的攻击方式进行了较为全面的分类总结,在此基础上对目前主要的反钓鱼方法进行了分类研究,并对各种方法的优缺点进行了相应的分析;提出网络钓鱼在三个方面的发展趋势,并得出结论:只有将各种防御方法很好地结合起来才能更好地应对手段不断翻新的网络钓鱼攻击。

关键词:网络钓鱼;钓鱼邮件;反钓鱼
中图分类号:TP393.08 **文献标识码**:A **文章编号**:1673-629X(2008)09-0115-04

Survey on Anti-Phishing Approaches

NI Tian-hua, ZHU Cheng-rong
(Dept. of Computer Science and Technology, Tongji University, Shanghai 201804, China)

Abstract: Phishing is an online identity theft that aims to steal sensitive information. Due to the widespread using of Internet, the number of phishing attacks grows rapidly which results in a lot of damages. Researchers as well as the IT industry have identified the urgent need for anti-phishing solutions and a number of solutions to mitigate phishing attacks have been proposed. Firstly introduces the basic concept of phishing. Secondly, it catalogues the existing phishing schemes as well as the anti-phishing approaches. As to each approaches, analyses the advantages and disadvangtages. Finally, points out the trends of the phishing scams and the anti-phishing methods.

Key words: phishing; phishing-mail; anti-phishing

0 引 言

随着互联网技术及应用的飞速发展和普及,我国网民的数量急剧增长。据统计,国内约有4000万的个人网上银行用户,有八成以上习惯在网上进行交易或炒股。iResearch的最新数据显示,2007年一季度国内第三方支付市场交易额规模达到160亿元,比去年同期增长了4倍多。然而,网上交易在给人们带来方便的同时,也成了不法分子的攻击目标。2007年上半年,CNCERT/CC接收的1813件非扫描类网络安全事件中网络钓鱼事件数量最多,占所有接收事件的35%[1]。

巨大的经济利益使得钓鱼攻击层出不穷,攻击手段不断翻新,导致众多互联网用户对网络信息安全缺乏信心,对于互联网进一步的健康发展产生了严重的影响。

收稿日期:2007-12-18
作者简介:倪天华(1983-),女,江苏通州人,硕士研究生,研究方向为容错计算与信息安全;朱程荣,副教授,硕士生导师,研究方向为容错计算与信息安全。

1 网络钓鱼的含义

网络钓鱼主要是指利用互联网进行的一种欺诈行为。它通过诱骗用户提供其个人账户和密码、信用卡信息、社保编号等个人资料,获得用户的某种身份信息,进而窃取用户的个人财产[2]。

由于这种攻击是利用欺骗手段让用户自己"上钩",主动提供个人信息的,所以形象地称其为"钓鱼"。网络钓鱼(Phishing)一词起源于20世纪90年代中期,鉴于早期的攻击者是用电话线作案,所以他们常常用Ph(one)来取代fishing中的f,就形成了今天的Phishing一词。

起初,攻击者利用电子邮件作为诱饵,盗用美国在线用户的账号和密码。早期的网络钓鱼绝大多数只是对虚拟财富的窃取,如QQ号码、网游账号及装备等。但随着网上银行和网上交易的普及,"经济利益"成为犯罪分子不断追求技术突破的源动力。网络钓鱼的手段不断翻新,并逐渐与其他技术手段相结合,但目标始终是个人身份标识数据,途径都是通过计算机网络窃取。因此,广义上网络钓鱼用来指"在线身份窃取"(Online Identity Theft)这一类攻击。

2 网络钓鱼的分类

由于网络钓鱼手段纷繁复杂,对其的分类也是多种多样。总体上,钓鱼实施过程可以分为诱骗阶段和信息获取阶段。

2.1 诱骗阶段

(1)欺诈邮件。

这是早期的攻击者常用的攻击手段。他们利用SMTP协议不验证发送人身份的漏洞伪装成著名金融机构、银行和在线交易网站向用户发送欺诈邮件。邮件中声称用户的密码被多次误入,怀疑有人试图盗窃用户账户;或者声称网站向用户提供升级服务,诱骗用户通过邮件中提供的超链接到网站去确认或修改账户和密码。而这个超链接所指向的网站是攻击者精心设计的仿冒网站,用户一旦输入个人信息,这些信息就会通过各种途径发送给攻击者,从而被其利用造成经济上的损失。

(2)即时通讯软件。

即时通讯软件是类似于电子邮件的互联网通信方式,与电子邮件最大的区别就是即时性。常见的即时通讯软件包括:OICQ、MSN、网易泡泡、淘宝旺旺等。当用户在使用即时通讯软件时,诈骗者就使用虚假信息欺骗用户,从而获取隐秘的个人信息。

(3)搜索引擎。

随着Web2.0时代的来临,以Google、百度为首的搜索引擎为普通用户带来方便的同时,也给不法分子提供了可乘之机。利用普通用户上网的特点及自动搜索算法的规则使得精心设计的钓鱼网站排列于搜索结果的显著位置,用户一旦点击,便落入陷阱。

(4)恶意引导。

攻击者利用DNS欺骗、修改HOSTS文件、CSS跨站脚本语言等方式将用户由正常的域名引向其精心制作的虚假网站,从而获取用户信息。

(5)恶意代码植入。

攻击者通过恶意网站和浏览器漏洞等系统漏洞在用户计算机中植入恶意程序进行信息窃取。其中,前两种是用的社会工程学方法,后两种则主要是应用黑客技术。

2.2 信息获取阶段

(1)假冒网站。

钓鱼者首先建立域名和网页内容都与真正的网上银行系统、网上证券交易平台极为相似的网站。通过诱骗阶段的各种方式将用户引入事先建立的虚假网站,进而获取用户的账号、密码等信息。根据APWG的最新统计,高达95.2%的网络欺诈是针对金融机构的[3],最常被仿冒的前三家公司为:Citibank、eBay和Paypal。

(2)虚假的弹出窗口。

一些攻击者会将用户转移到真实的网址,但是转移之前制造假冒的弹出窗口,提示用户进行个人登陆的操作,一些普通用户会误认为是真实网站的一部分而进一步按照提示操作,最后泄露了个人身份信息。

(3)恶意代码。

通过事先植入的恶意代码,当用户进行网上交易时通过键盘记录、屏幕截取等方式记录用户个人信息,通过网络发送给攻击者。

3 网络钓鱼的防御方法分类

3.1 基于邮件服务器

由于大量的钓鱼攻击是通过欺诈性的电子邮件实现的,因此比较直接的防御技术就是对钓鱼邮件进行过滤。

(1)针对SMTP协议漏洞。

SMTP是现在通用的邮件传输协议,邮件服务器处理邮件的步骤是由SMTP协议来规定的。SMTP首先要求进行TCP对话,这时接收方的邮件服务器收到发信方的IP地址,然后通知发件人的邮件地址和收件人的邮件地址,最后再发送包括邮件头在内的邮件数据。接收的邮件包括三项发件人信息:发信方的IP地址、发信方通知的发件人邮件地址和邮件头部内记述的发件人地址。但这三项信息中,只有发信方IP地址是确凿可靠的,另外两项信息由于发件人邮箱或服务器的设置问题,总是可以伪造的。

针对协议的这一漏洞,可以人为地增加对发件人身份的认证机制,进而将钓鱼邮件用类似于过滤垃圾邮件的方法屏蔽掉。

文献[4]提出了一种SEFAP(Signed Email for Anti-Phishing)系统模型。SEFAP系统由服务器和客户端两部分组成。其中,SEFAP服务器位于邮件服务器端,由系统管理器、私钥生成器(PKG)、提取器、签名生成器、签名验证器、同步器和调度器组成;SEFAP客户端位于发送者的机器上,主要负责在用户发送邮件时的签名工作。在邮件系统初始化时,SEFAP服务器运行私钥生成器来生成系统参数,并发布公共参数。进而运行提取器为每个新注册的用户生成各自的私钥。当用户发送邮件时,SEFAP调用签名生成器对其进行签名。当用户请求阅读接收到的邮件时,SEFAP使用签名验证器验证并输出一布尔值指示邮件服务器的动作。如果值为真,SEFAP通知邮件服务器像原邮件系统那样处理该邮件;否则,该邮件被挂起,并采取相应的措施如阻止接收或给用户相应的告警。其中私钥生

成器中所使用的签名机制由系统管理器指定,并可以更改。

文献[5]则是通过对原始邮件建立发件人的可信度参数,利用相应规则进行过滤。

(2)针对伪URL地址。

根据伪URL的特点大体可以分为以下几类:

(1)可见链接与实际链接不同;

(2)使用点分十进制IP地址而非域名;

(3)对超链接用特殊字符进行恶意编码;

(4)使用与目标极相似的假冒域名。

文献[6]通过LinkGuard算法对URL进行检测,属情况(1)(4)则判定邮件为钓鱼邮件,属于(2)(3)则将其归类为可能的钓鱼邮件。将此系统应用于服务器端则可对钓鱼邮件有一定的防范作用。但是必须指出,此方法存在漏判和误判的情况。

(3)基于模式识别的。

由于假冒网站在风格上模仿真实站点,因此在视觉上有很大的相似性,这也是大量普通网民被骗的原因。针对这一特点,文献[7]提出了从视觉相似这个角度进行反钓鱼的方法。

首先,将需要保护的合法网站注册在反钓鱼数据库服务器中。事先对合法网站进行特征提取,将其存入反钓鱼数据库。其次,设置反钓鱼代理。代理中基于EMD算法[8]对邮件中的链接的网站进行相似度计算和分类,如果相似度超出一定的域值则进行钓鱼网站的报告。在实验阶段,对于报告的钓鱼网站进行人工的复查,以对分类规则进行改进。将该系统应用于邮件服务器上,可以对钓鱼邮件进行监控。

3.2 基于浏览器

(1)基于脚本语言。

获取密码的过程大多是通过网页脚本语言(JavaScript, VBScript等)的强大功能实现的,于是最直接的防御方法是关闭对脚本语言的支持。然而这是不现实的,因为大部分正常的网站也应用了大量的脚本语言来完成相应的功能。

AntiPhish则是对JavaScript进行了相应的控制。当网页的焦点是在文本输入的时候,禁用脚本语言,从而防止输入信息的同时,攻击者利用脚本语言进行信息的窃取;而当焦点不再是信息输入的时候,再激活对脚本支持[9]。AntiPhish主要针对的是利用脚本语言进行键盘、屏幕记录的攻击,是应用于Mozilla FireFox的浏览器插件。然而在扩展到IE浏览器时由于Linux和Windows的系统机制不完全相同,使得这种特定的方式在IE上失效[10]。

(2)基于黑名单。

利用数据库中存储的Phishing黑名单对站点进行检查,如果用户访问的站点在此黑名单之列,就向用户发出警告。此类应用包括IE7.0、EarthLink的ScamBlock[11]、PhishGuard[12]、Netcraft[13]以及Google Safe Browsing on Firefox[14]等等。其中的黑名单通过用户举报、蜜网监测等方式不断更新。这与传统防病毒软件利用特征码反病毒类似。

但是,据APWG最新报告显示,钓鱼网站的平均存活时间为3.8天,而最长的也只有30天[3],因此,黑名单也具有一定的局限性,并且它无法预防新的钓鱼攻击。

3.3 基于协议的改进

SSL协议位于应用层之下传输层之上,在SSL安全机制中,客户端首先与服务器建立连接,服务器把它的数字证书和公钥一并发送给客户端,客户端随机生成会话密钥,使用从服务器得到的公钥对会话密钥进行加密,并将会话密钥在网络上传递给服务器,而会话密钥只有在服务器端用私钥才能解密。

传统SSL协议只保证会话的机密性,由于证书机构(CA)可能错发证书,因此不能保证基于SSL协议的会话一定是与合法站点进行的。只有当进一步检查证书内容时才能得出结论,而这一点是一般用户不可能做到的。

文献[15]提出了一种基于ID的SSL协议。该协议利用基于双线性映射对的公钥机制实现客户端与服务器端之间的通信。每一个合法站点在私钥生成机构认证一次,由私钥生成机构为其分配私钥。用户第一次进入该站点时,对服务器提供相应的公共参数进行存储。之后每次与该服务器建立加密会话时都使用该系列公共参数。由于服务器站点在私钥生成机构注册了唯一的身份,因此使用相应的公共参数只能与对应的站点建立加密会话。并且由于私钥生成机构只对站点做一次认证,钓鱼站点无法对合法站点进行身份伪装。

该方法虽然很好地解决了服务器的认证问题,但是没有解决直接钓鱼站点而非仿冒的攻击。

3.4 小 结

文中仅仅介绍了反钓鱼方法的系统架构,并没有给出具体实现细节。

其中,3.1节中的基于邮件服务器的方法主要是针对欺诈邮件,把钓鱼攻击阻止在诱骗阶段;3.2节中基于浏览器的方法主要是针对仿冒网站等信息获取阶段攻击的防御;3.3节则是从较底层的架构出发给出的整体解决方案;对于恶意代码等的攻击则需要将上述方法与反恶意代码技术相结合,通过反病毒软件检

测非正常的程序行为以阻止钓鱼攻击。

4 结束语

目前,“网络钓鱼”的发展趋势呈现三个方面特点,即多种手段结合、隐蔽性、多样性。多种手段结合是指“网络钓鱼”由单纯的社会工程学的欺骗演变为综合了黑客技术、漏洞注入、脚本欺骗等手段的网络犯罪行为。隐蔽性是指利用各种黑客技术,使“网络钓鱼”行为更加隐蔽,用户更加难以察觉。多样性是指“网络钓鱼”的目标由单一的网络银行转化为各种在线销售系统、电子商务网站及股票交易等在线系统。随着网络钓鱼的手段不断翻新,只有将各种防御方法相结合,才能更加有效地防止在线身份窃取。

参考文献:

[1] CNCERT/CC2007 年上半年网络安全工作报告[EB/OL]. 2007. http://www.cert.org.cn/articles/docs/common/2007082123431.shtml.

[2] 杜跃进. 在线身份窃取攻击[J]. 网络安全技术与应用, 2005,8:7-9.

[3] Phishing Activity Trends Report for the Month of June,2007 [EB/OL]. 2007-06. http://www.antiphishing.org/ .

[4] Ren Qiong, Mu Yi, Susilo W. SEFAP: An Email System for Anti-Phishing[C]//Proceedings of the 6th IEEE/ACIS International Conf. on Computer and Information Science. Australia: IEEE Press, 2007:782-787.

[5] Inomata A. A Novel Mail Filtering Method Again Phishing [C]//Pacific Rim Conf. on Communications, Com-puters and signal. [s.l.]: IEEE Press, 2005:221-224.

[6] Chen Juan, Guo Chuanxiong. Online Detection and Prevention of Phishing [C]//Proceedings of the first International Conf. on Communications and Networking. China: IEEE Press, 2006:1-7.

[7] Liu Wenyin. An antiphishing strategy based on visual similarity assessment[J]. Internet Computing, 2006, 10(2):58-65.

[8] Fu A Y. Detecting Phishing Web Pages with Visual Similarity Assessment Based on Earth Mover's[J]. IEEE Trans. on Dependable and Secure Computing, 2006, 3(4):301-311.

[9] Kirda E, Kruegel C. Protecting users against phishing attacks with AntiPhish[C]//Proceedings of 29th Annual International Conf. on Computer Software and Applications. [s.l.]: IEEE Press, 2005:517-524.

[10] Raffetseder T. Building Anti-Phishing Browser Plug-Ins: An Experience Report[C]//Proceedings of the 3rd International Workshop on Software Engineering for Secure Systems (SESS'07). Austria: [s.n.], 2007:1-7.

[11] EarthLink. ScamBlocker[CP/OL]. 2004-07. http://www.earthlink.net/ software/free/toolbar/.

[12] PhishGuard.com. Protect Against Internet Phishing Scams [CP/OL]. 2003-01. http://www.phishguard.com/.

[13] Netcraft. Netcraft toolbar[CP/OL]. 2004-07. http://toolbar.netcraft.com/.

[14] Google Safe Browsing for Firefox[CP/OL]. 2005-10. http://www.google.com/tools/firefox/safebrowsing/.

[15] Tan Chik How, Teo Joseph Chee Ming. Protection Against Web-based Password Phishing[C]//International Conference on Information Technology (ITNG'07). [s.l.]: IEEE Press, 2007:754-759.

(上接第 111 页)

Computer Vision, 1996, 19(2): 129-146.

[7] 邓亚峰,苏光大,傅 博. 一种基于 AdaBoost 的快速动态人脸检测算法[J]. 计算机工程,2006,32(11):222-224.

[8] Yin Jian-qin, Li Jin-ping, Han Yan-bin, et al. A New Color-Based Face Detection and Location Method by Using Support Vector Machine[C]//In Proc. IEEE Conf. Control, Automation, Robotics & Vision. Kunming, China: [s.n.], 2004: 838-841.

[9] Han Yan-bin, Liu Ming-jun, Li Jin-ping. Face Detection and Location Based on Skin-color Modeling and Geometrical Features[J]. Computer Science, 2006, 33(s):311-313.

[10] 陈锻生,刘政凯. 肤色检测技术综述[J]. 计算机学报, 2006,29(2):194-207.

[11] 李士进,朱跃龙,王志坚. 基于多分类器组合的多角度彩色人脸图像检测[J]. 小型微型计算机系统,2004,25(8): 1506-1509.

[12] 李国辉,梅魁志,袁泽剑. 基于肤色模型和贝叶斯判别的人脸检测[J]. 计算机应用, 2006,26(8):1854-1857.

[13] 艾海舟,梁路宏,徐光祐,等. 基于肤色和模板的人脸检测[J]. 软件学报,2001,12(12):1784-1792.

[14] 王文宁,王汇源,常新华. 一种新的基于对称特征的彩色人脸定位方法[J]. 计算机工程与科学, 2006, 28(10):54-56.

[15] 陈启泉,邱文宇,陈维斌. 标准正面人脸图像的特征提取[J]. 华侨大学学报:自然科学版,2000,21(4):413-418.

[16] Hsu R L. Face Detection in Color Images[J]. Pattern Analysis and Machine Intelligence, 2002, 24(5): 696-706.

[17] 吕东辉,王 滨. YCbCr 空间中一种基于贝叶斯判决的肤色检测方法[J]. 中国图象图形学报,2006,11(1):47-52.

[18] 刘 洁,张汉灵. 一种新的基于肤色模型的人脸检测算法[J]. 计算机工程与应用,2006(11):70-72.