第32卷 第14期 计 算 机 工 程 2006年7月
Vol.32 № 14 Computer Engineering July 2006
·安全技术· 文章编号：1000—3428(2006)14—0125—02 文献标识码：A 中图分类号：TP393.08

Phishing攻击行为及其防御模型研究

张 博，李伟华
(西北工业大学计算机学院，西安 710002)

摘 要：仿冒(Phishing)危害愈演愈烈，针对其攻击行为进行了详细的分析与介绍，其中使用了建立攻击森林和对攻击进行分类等方法，进而建立了Phishing攻击模型。提出了相应的Phishing攻击的防范理论体系和具体措施。同时高起点地分析了IPv6环境下的Phishing攻击及其防御。

关键词：仿冒；IPv6；攻击行为；防御体系

Study on Phishing Attack Behaviours and Defence Model

ZHANG Bo, LI Weihua
(College of Computer Science, Northwestern Polytechnical University, Xi'an 710072)

【Abstract】 The phishing hazard becomes more and more terrible. This paper introduces and analyses the phishing attack behaviors. It adopts the methods of building the attack forest and sorting the attack behaviours and then constitutes the attack model. It presents the anti-phishing theory system and some countermeasure. Furthermore it analyses the phishing attack and defence based on the IPv6.

【Key words】 Phishing; IPv6; Attack behaviors; Defense system

1 概述

仿冒(Phishing)是指攻击者利用欺骗性的电子邮件和伪造的Web站点来进行诈骗活动。受骗者往往会泄露自己的个人信息和财务数据，包括个人的技术资料、联系方式、E-mail、银行卡号、账户、密码等。反仿冒工作组(APWG)在其网站上报道，2004年6月Phisher共发动1 422次新攻击，比5月份的1 197次攻击多19%，比2003年12月份的116次攻击多12倍以上。phishing正在以平均每月38%的速率递增，同时假冒网站的数量正在以每月24%的数量增长[1]。

在我国，仿冒已有案例，例如，2005年1月14日中央电视台经济信息联播报道了假网站冒充工商行诈骗80多万元的事实。而据最新的APWG的统计，中国(包括香港、台湾)的仿冒攻击已经超越其他国家，仅次于美国位居第2。随着仿冒手段的不断发展和翻新，我国金融领域也必将出现更多的仿冒现象，研究仿冒防御，对填补国内空白以及金融银行领域免遭巨大损失具有重要的意义。

国外对phishing的研究从2004年以来逐渐增多，但还没有形成成熟的理论体系。主要是phishing的起源、概念理解，以及phishing的攻击过程和初步的防御方法。国内相关phishing的报告从2004年下半年开始出现，主要是phishing对行业造成危害的新闻，以及phishing的概念普及。

2 关键技术研究

针对当前的phishing现状，在实际的研究当中，为了更好地防御phishing，本文进行了如下的研究。

2.1 phishing攻击行为的研究

对于当前的各种phishing攻击，从攻击的角度来分析攻击者的攻击意图、攻击手段，对现有的各种phishing攻击进行分类，并对每种攻击进行分别详尽的分析。解决的关键技术包括：一次成功地phishing攻击的各个阶段的合理划分；phishing行为的分析；phishing攻击的有效分类；各种攻击树的建立(包括一般攻击树的建立和分类攻击树的建立)。为了对phishing攻击行为做到有效分析，首先对要分析的原始事件作清楚、明确地描述。为此对收集到的phishing事件采用统一的时间格式，抽象为一个向量组。而且复杂的攻击通常总是按照一定的步骤和路径进行的，即同一复杂事件是因果相关的。所以在研究当中采用了事件的相关性分析：

(1)运用管理分析技术发现每条审计记录内部不同属性之间的相互依赖的模式；

(2)通过序列分析从时间角度在事件序列中寻找事件间发生的潜在模式，在此基础上定义了典型的前后相关的phishing攻击。

对phishing攻击进行了合理的规范和分类，采用基于具体应用的多维分类方法：一维为攻击工具；另一维为上下文环境，再一维是攻击方法。以关联与序列模型推理的基本思想建立攻击树或者森林。即以关联规则为基础建立攻击描述森林，然后将不同对象的攻击序列模式叠加在攻击描述森林上，最后得到一个全局的攻击模式。而且将不同层次的数据汇总，用以发现更为复杂的攻击模式。这些通过改进的SWF算法来实现。

2.2 phishing攻击模型研究

phishing攻击模型是同类攻击行为的统一和简洁的描述，为防御机制建立打下了良好的基础。利用图论和必要的数学工具，建立起相应数学模型，生成攻击模型图。解决的关键技术包括：对模型的定义方法；图形模型中各个点、线所代表含义；复杂攻击模型的提炼；模型中如何加入经济、人为等因素；采用统一的理论模型对各种攻击行为模型化。

基金项目：国家"863"计划基金资助项目(2003AA142060)；国家网络与安全管理中心项目(2004-研1-917-C-020)
作者简介：张 博(1977—)，男，博士生，主研方向：网络安全，多媒体技术；李伟华，教授、博导
收稿日期：2005-09-22 **E-mail：**zhangbo@mail.nwpu.edu.cn

建立一种能够描述各种phishing攻击的模型图。建立模型图的方式为[2]：用节点来表示信息或者访问权；用相连的边来表示获得信息或者访问权的手段和方法，边还表示可能性、花费等，这样能够通过图形模型来量化phishing，并能对攻击进行经济分析。实际研究中基于图形的模型是分为一般和特殊两个方面来描述的。

举一个最简单的例子来说明所要研究的模型的架构，如在一次银行账户的简单攻击中，它的最简单的模型表示如图 1 所示。

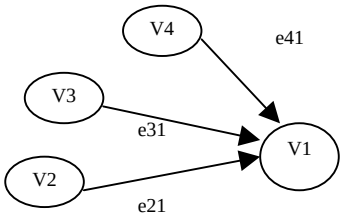

图 1 一次简单的代表依次银行攻击的图形模型

顶点 V1 代表受害者的银行账户，V2 代表受害人的银行存款信息，e21 就是 phisher 猜测受害人信息的过程；V3 代表受害人的亲戚关系，e31 就是受害人与其它的亲戚账户的关联性；V4 代表受害者账户的一次访问，e41 就代表执行支付的相应的行为。

进一步如果知道受害者银行账号信息或者其他的因素，模型图就能基本反映现实情况。

2.3 phishing 的防范理论与体系研究

形成可行的减少和避免被 phishing 攻击的可靠性的理论体系，主动对 phishing 进行预警。充分调动各种网络安全技术手段，综合性地对 phishing 攻击进行防范。在现有的安全条件下最大限度地保证系统免受攻击。解决的关键技术包括：全面的防御体系；各种现有网络安全防御手段的结合方法；防御方法中的有效的通信机制；构建完整的 phishing 防范的理论。

现今的phishing方法逐渐增多，没有万能的方法来抗击所有的攻击。所以使用现有的信息安全技术和手法的结合，来防范目前的乃至未来的phishing攻击是非常切合实际的。为了达到最好的保护，采用客户端、服务器端、组织级三层配置的安全防御体系[3]。主要采用分布式技术、JINI技术建立起有效防御方法的通信机制综合管理。客户端：包括了使用者的个人电脑；服务器端：包括了商业的互联网系统和客户的应用；组织级：分布式技术和第三方管理服务。而且在多层次的防御方法中建立了有效的通信机制，这主要通过各级的代理来实现。

在客户端，因为现在安全意识的逐渐加强，桌面系统配置成使用多个桌面保护代理，安装反病毒软件和网络防火墙。通过各种安全技术的联合使用，添加上 phishing 常用攻击手段中的基本特征，各种桌面代理共同发挥作用，减少了 phishing 攻击的发生。

在服务器端，反phishing与具体的Web安全应用相结合，根据可信任区域(Trusted Credential Area，TCA)的概念[5]，设计实现一个收集可信任信息的信任收集模型。信任收集器接受浏览器 3 个方面的输入，用来确认页面公共密钥(PK)；指出页面额外认证和公钥的URL，主要通过<META>标签来实现；在SSL握手阶段服务器提供的认证。一个建立TCA可信任区域的模块如图 2 所示。

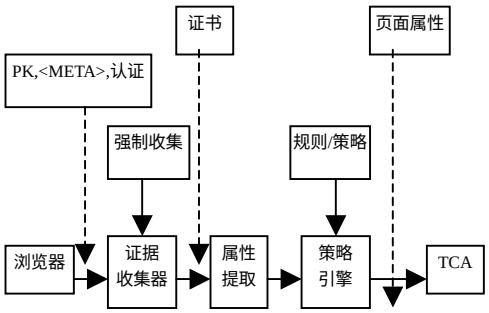

图 2 建立 TCA 可信任区域的模块

组织级是在保护客户端和服务器端的基础上，对 phishing 的深度防御机制。包括自动的邮件发送地址确认、电子邮件的数字签名、对域名进行监测等来配合实现。

2.4 IPv6 协议下的 phishing 特殊安全防范研究

IPv6 协议对网络安全提供了更有力的支持，但也存在安全问题，本着前瞻意识，充分发挥该协议网络安全方面的优势，同时解决存在的安全问题，进而高起点地研究新协议下的防 phishing 技术，对于 phishing 在 IPv6 协议下的具体防范，做好了理论上的支持。解决的关键技术包括：IPv6 本身以及在过渡阶段存在的安全问题；IPv6 协议下的简单攻击；IPv6 环境下的 phishing 防御基础研究。

分析 IPv6 协议本身的特点，在 IPv4 基础上对 IPv6 进行对比分析。利用现有的防火墙和 IDS 等网络安全技术软件对 IPv6 的安全性进行分析，配合 IPv6 实验网络，总结出 IPv6 协议的安全性。IPv6 扩展了地址空间，协议本身提供加密和认证功能，因此，面向 IPv6 的 phishing 防御着重解决：(1)大规模网络环境下的 phishing 攻击行为。由于 IPv6 支持超大规模的网络环境，面向 IPv6 的 phishing 攻击的研究要解决大数据量的问题，需要融合分布式结构和高性能计算技术。(2)认证和加密情况下的网络监听。IPv6 协议本身支持加密和认证的特点，极大地增加了面向 IPv6 的 phishing 攻击行为监听网络数据包的难度。

在 IPv6 环境下，根据 IPv4 环境下的防范理论，结合 IPv6 协议本身特点，以及模拟的 phishing 攻击，总结出防范 Phishing 主要方法和手段。引入计算机免疫学的相关理论，对于前期在 IPv4 环境下的研究的攻击模式、攻击特点等理论进行“细胞记忆”。提供相关的疑似特征，利用免疫学中的自体与非自体、亲和力、克隆选择、学习机制和免疫记忆的相关概念，结合 IPv4 环境下的防范理论建立 IPv6 协议下的 phishing 的特殊防范理论。

2.5 phishing 防范的具体措施研究

针对具体的 phishing 攻击，编制了相对应的防范软件。利用现有的技术优势，使 phishing 防范由被动变为主动，弥补以前各种网络安全手段在 phishing 防范上面的不足。解决的关键技术包括：防御机制的代码实现；防御体系中的通信机制的代码实现；IPv6 环境下的 phishing 防御的代码研制；可信任区域(Trusted Credential Area，TCA)理论模型的代码实现。根据不同分类的 phishing 攻击写出了详尽的分析报告和应对措施报告。对系统的 phishing 攻击的防御提出有效的模型和理论。在此基础上，编制相应的 phishing 防御软件，通过前期的各种数据和图表建立相应的知识库、模型库，应用决策支持系统(DSS)对知识库的表示、获取以及推理机构方面的优势，采用分布式知识存储、分布式并行推理、分布式

（下转第 135 页）