计算机研究与发展
Journal of Computer Research and Development
DOI:10.7544/issn1000-1239.2021.20210855
58(10):2120-2139,2021

网络犯罪的检测分析技术

洪 赓 杨 森 叶 瀚 杨哲慜 杨 珉
(复旦大学计算机科学技术学院 上海 200433)
(ghong17@fudan.edu.cn)

Detection and Analysis Technology of Cybercrime

Hong Geng, Yang Sen, Ye Han, Yang Zhemin, and Yang Min
(*School of Computer Science*, *Fudan University*, *Shanghai* 200433)

Abstract With the rapid growth of information technology, people's daily activities have been gradually moving to cyberspace. Online activities also play an increasingly important role national economy. While the Internet greatly facilitated our daily life, more and more criminal activities that threaten our daily life, have also moved to cyberspace. Therefore, how to understand, evaluate, prevent, and combat cybercrimes have become the focus of attention of academia, industry, and law enforcement agencies. Recently, researchers pay much attention to the prevention, evaluation, and countermeasures of cybercrimes. However, until now, only a few researchers focus on the overview of cybercrime. Also, there is an urgent need for systemization of the entire cybercrime kill chain. This paper starts from some classic cybercrime attacks such as phishing, scam, and cryptojacking, and then an in-depth analysis of their supporting techniques is conducted, including blackhat SEO and typosquatting. To analyze the cybercrime kill chain, we also investigate the cybercrime infrastructures such as underground market, botnet, and money laundering. Finally, we discuss the existing challenges and trends of cybercrime research.

Key words cybercrime; detection and analysis technology; phishing; network scam; cryptojacking

摘 要 随着信息技术的高速发展,越来越多的生产生活逐渐转移到网络空间进行,国民经济对网络空间的依赖也日益凸显.互联网带来便利的同时,越来越多的犯罪从传统线下转移到网络空间中进行,威胁人民群众的日常生活安全.因此,如何理解、评估、预防、打击网络犯罪,成为学术界、工业界和相关执法部门的关注重点.近年来,研究人员持续关注各种网络犯罪及对应的防范、评估、反制技术.但目前针对网络犯罪总体综述研究较少,亟需对网络犯罪产业链组成部分进行全面且详细的梳理.将以钓鱼(phishing)、诈骗(scam)、恶意挖矿(cryptojacking)等经典网络犯罪攻击方式为切入点,深入分析包括黑帽搜索引擎优化(Blackhat SEO)、误植域名(typosquatting)在内的相关支撑技术,详细揭露地下市场(underground market)、僵尸网络(Botnet)和洗钱渠道(money laundering)等网络犯罪基础设施,剖析网络犯罪产业链,最后讨论了网络犯罪研究中仍存在的挑战,并展望未来研究方向.

关键词 网络犯罪;检测分析技术;钓鱼攻击;网络诈骗;恶意挖矿

中图法分类号 TP391

收稿日期:2021-08-23;**修回日期**:2021-08-30

基金项目:国家自然科学基金项目(U1836213)

This work was supported by the National Natural Science Foundation of China (U1836213).

通信作者:杨珉(m_yang@fudan.edu.cn)

网络犯罪,是指使用计算机、网络或相关硬件设备促成或实施的任何犯罪[1].广义上,世界上第1次记载的网络相关犯罪可以追溯至1834年2名劫匪侵入法国电报系统并从股票市场窃取信息[2].伴随着移动生态的快速发展,人民的生活数字化程度逐渐提高,越来越多的生产生活逐渐转移到网络空间进行,网络犯罪也随之更加频繁发生,如2016年美国56名嫌疑人冒充国税局或移民工作人员进行电信网络诈骗,导致1.5万人上当受骗,涉案金额超过3亿美元[3].近年来,国民生活逐步从线下转移到线上进行,网络犯罪也随之愈演愈烈:2019年仅在俄罗斯就发生了一万多起新型网络犯罪[4].同时,英国的研究人员发现新冠疫情出现以来网络诈骗事件数量持续增长,电信网络诈骗者正在利用疫情对受害者实施诈骗[5].

近年来,研究人员主要关注各类网络犯罪的防范、评估、反制技术的发展.但目前针对网络犯罪总体综述研究较少,缺乏对网络犯罪产业链进行全面梳理.

如图1所示,网络犯罪产业链庞大而复杂,其中涉及到攻击形式、支撑技术、基础设施的协调配合.网络犯罪攻击形式是网络犯罪产业链中直接接触终端受害者的部分,其最易被人们感知,也直接造成人民群众财产经济损失.然而,具体的攻击形式是网络犯罪产业链的表层体现,在相关犯罪行为背后,通常是由网络犯罪支撑技术为其提供充足的技术支持.此外,网络犯罪攻击形式和支撑技术均严重依赖于交易平台、网络设施、洗钱渠道等网络犯罪相关基础设施,它们为支撑技术的成功应用和网络犯罪攻击的成功实施提供相应的基础服务保障.

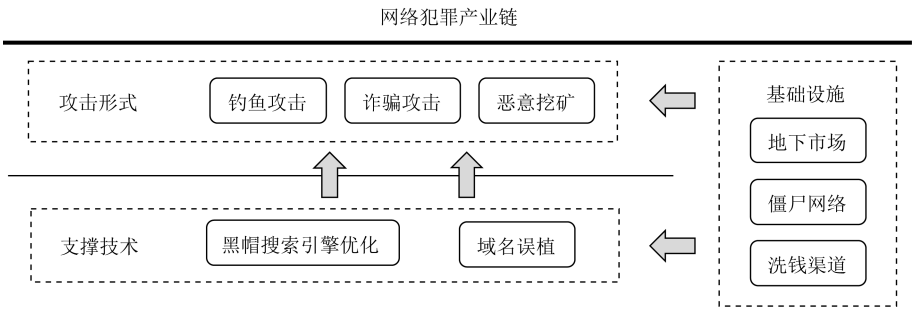

Fig. 1　Illustration of cybercrime industrial chain

图1　网络犯罪产业链示意图

为此,本文将以钓鱼(phishing)、诈骗(scam)、恶意挖矿(cryptojacking)等经典网络犯罪形式为切入点,介绍其犯罪特点和逻辑,进而对以黑帽搜索引擎优化(blackhat search engine optimization, Blackhat SEO)、误植域名(typosquatting)为代表的相关支撑技术进行梳理,最后对网络犯罪所赖以生存交易平台、网络设施、洗钱渠道等基础设施(如地下论坛(underground forum)、僵尸网络(Botnet)和洗钱渠道(money laundering)等)进行解析,串联网络犯罪生态的各个组成部分,剖析网络犯罪产业链.由于篇幅所限,本文主要关注于网络犯罪产业链中环节的国内外重要学术研究成果,着重介绍上述网络犯罪所涉及的概念,梳理其中检测分析技术和存在问题,最后讨论仍存在的挑战并展望未来研究方向.

1　网络犯罪的攻击形式

网络犯罪的攻击形式是网络犯罪庞大产业链的具体呈现.其在整个网络犯罪产业链中负责直接接触和攻击受害者,并产生经济利益,为产业链的发展提供源源不断的资金支持.为了迷惑受害者,网络犯罪的攻击形式通常还会和社会工程学、人机交互设计相结合,以达到最佳的攻击效果.本节将以钓鱼攻击、网络诈骗和恶意挖矿3种经典的网络犯罪的攻击方法为例,讨论和梳理相关研究工作和技术趋势.

1.1　钓鱼攻击

钓鱼攻击(phishing),指通过伪装成权威的网站或者机构来获取受害者用户户名、密码和银行卡信息等个人敏感数据的犯罪过程.犯罪分子通常会声称自己来自网络银行如支付宝、PayPal,或者权威机关如公安机关、电信部门等获取用户的信任,再诱导受害者点击网站链接到外观与被仿冒的网站相差无几的假网站输入个人资料.在获得受害者的敏感数据后,犯罪分子可以通过出售等非法途径进一步牟利.

作为一种经典的网络犯罪,钓鱼攻击的普遍性使得其一直是学术界和工业界研究的重点.早期的

钓鱼攻击往往场景较为单调,因此基于黑、白名单的检测方法被广泛应用.但随着对抗的不断深入,研究者的重心开始转向基于机器学习或者启发式算法等更高效准确的检测方法上.同时,为了防御将来可能发生的钓鱼攻击,研究人员通过对钓鱼攻击不同的环节进行深入分析,总结犯罪分子的技术特点和表现特征,以提出更好的防御和对抗方法.本节将从钓鱼攻击的检测方法、防御技术、对抗方法和生命周期特点等维度来总结近年来学术界针对钓鱼攻击的研究成果.

一般而言,钓鱼网站获取受害者信任最直接的方式就是保持同原网站在视觉上高度的相似性,以达到以假乱真、鱼目混珠的效果.但这种视觉上的相似同样可以被研究人员用来检测钓鱼网站.考虑到对网站截图进行直接对比不仅效率较差,而且准确度更低,Lin 等人提出了基于深度学习的检测工具 Phishpedia[6],通过定制的目标检测模型提取页面中的品牌图标信息,然后利用基于迁移学习的孪生神经网络进行检测.针对传统图像相似检测方法无法检测未知图形的问题,Abdelnabi 等人创新性地使用了三元卷积神经网络模型,基于网站屏幕截图,实现了基于相似度检测钓鱼攻击的方法 VisualPhishNet[7].

除视觉相似度等直观因素的影响外,HTML 代码的相似度同样可以作为检测钓鱼攻击的重要依据.Cui 等人从 HTML 代码内的标签入手,通过统计不同标签在网页内出现的频率将不同的 HTML 编码为等长的向量,然后通过定义比例距离的方式来判断不同网页的相似度.Cui 等人[8]将具有高相似度的网页用启发式的算法进行聚类,由此得到来自同一个模板生成的不同钓鱼网站集群.为了精准地获得网页中的文字内容,Tian 等人引入了视觉分析和光学字符识别等功能来解决攻击者对网站内容的混淆[9].

基于网站内容相似度的匹配技术也被 Yoon 等人用于检测暗网上的钓鱼网站[10].之前的研究大多着眼针对公开网站或机构的钓鱼,作为一类允许服务提供商和访问者使用匿名网络服务来隐藏身份的平台,暗网上是否存在着钓鱼攻击对研究者来说是一个非常有趣的问题.Yoon 等人证实了暗网中普遍存在钓鱼网站.

此外,犯罪分子在实施钓鱼攻击中的技术特点同样可以用来检测钓鱼网站.为了尽可能地模仿目标网站同时也为了降低成本,犯罪分子一般会在钓鱼网站中直接使用原网站的静态资源.根据这一特点,Oest 等人开发出了框架 Golden Hour,通过分析静态资源文件请求头的信息来检测模仿 PayPal 的钓鱼网站[11].为了进一步提高收益,犯罪分子通常会使用伪装技术来规避各大反钓鱼系统的检测,即通过识别网站访问者的身份来显示不同的页面.为了解决这一难题,Zhang 等人提出了基于代码路径执行区别特点的工具 CrawlPhish,针对性地检测钓鱼网站在客户端侧的伪装技术[12].

除了有效地检测钓鱼网站外,研究人员的另一个研究重点在于检查现有防御机制的有效性.国际域名(internationalized domain names, IDN)于 2003 年引入并标准化,支持来自各种语言的 Unicode 字符,由于来自不同语言的不同字符可能存在相似,随之出现了为网络钓鱼而创建的同形异义 IDN.浏览器通常会实施规则来检测可能冒充合法域名的同形异义 IDN,一旦识别到,浏览器将不再显示 Unicode,而是显示其 Punycode 以提醒用户.Hu 等人系统地评估了针对同形异义 IDN 的浏览器级防御[13].他们通过自动化测试,验证了所有主流浏览器规则均存在可被规避的盲点.

黑名单是用户抵御网络钓鱼的又一道技术防线,但由于其本质上的被动性,在钓鱼网站被加入黑名单之前,用户并不会收到浏览器的警告,因此不少钓鱼网站会利用伪装技术来延迟黑名单爬虫的检测.Adam 团队首次研究了伪装对于浏览器黑名单有效性的影响.他们在 2019 年开发了一个可扩展测试框架 PhishFarm[14],用于测试反网络钓鱼实体和浏览器黑名单对攻击者伪装的抵抗力[13].研究发现,黑名单在流行的移动浏览器中没有按预期发挥作用,使得这些浏览器的用户更容易受到网络钓鱼攻击.在 2020 年,该团队又开发了能持续检测并衡量钓鱼网站整体生态的框架 PhishTime,用以评估黑名单的性能[15].

除了防御手段外,研究人员还会主动部署反钓鱼系统以检测在野的钓鱼网站.反钓鱼系统在取得了较好结果的同时也催生出针对性的反分析技术,来对抗反钓鱼系统的检测.为了研究反分析技术的影响,Maroofi 等人针对谷歌重验证码、警报框和基于会话的规避等 3 种基于人工验证的反分析技术进行了测试实验[16].作者部署反分析技术的钓鱼网站,然后向主要的服务器端反钓鱼系统举报了这些网站,并持续监控它们在黑名单中的出现情况.Tian 等人在利用机器学习分类网络钓鱼域名的同时对钓鱼网站

的规避行为进行了研究,他们发现,在1 175个钓鱼网站中,超过90%的网站至少在一个月内成功避开了黑名单检测[9].Oest等人则另辟蹊径,尝试从钓鱼攻击发起者的视角来理解反钓鱼系统[17].通过分析收集到的超过2 300个钓鱼攻击使用的组件,发现钓鱼攻击发起者会利用请求信息、IP信息、域名信息等来主动过滤和对抗来自反钓鱼系统的检测.

为了研究网络钓鱼攻击的生命周期,Oest等人利用框架Golden Hour被动地对网络钓鱼页面的流量进行观测.通过一年时间的长期观测,Golden Hour记录了除爬虫外的480万次访问钓鱼网页的行为,并用其剖析钓鱼攻击的生命周期[10].Han等人则通过部署存在漏洞的蜜罐服务器,对钓鱼攻击的整个生命周期进行观测,包括攻击者在蜜罐上安装和测试网络钓鱼页面,到受害者的访问和个人信息发生泄露[18].通过长达5个月的数据收集,他们对网络钓鱼攻击的生命周期进行了全面的评估,包括犯罪分子的行为、钓鱼攻击的机制以及反钓鱼系统更新的实时性.

与其他网络犯罪活动一样,网络钓鱼攻击者和研究人员的对抗从未停止.钓鱼网站不断地使用内容混淆、伪装技术等规避方法来隐藏自身,逃避黑名单和反钓鱼系统等的检测;而研究人员则不断地将包括计算机视觉领域在内的最新研究成果应用于检测钓鱼网站.随着二者对抗形式的不断升级,在钓鱼攻击不断升级自己的工具、规避检测的同时,对钓鱼网站规避方法的研究,特别是犯罪分子如何逃避、对抗反钓鱼系统的检测也会进一步成为未来钓鱼攻击研究的重点.

1.2　诈骗攻击

随着互联网的不断加速发展,诈骗作为历史悠久的犯罪形式,衍生出了许多基于互联网中不同设施的诈骗形式.除了普通民众,电信运营商、零售企业、广告商都有可能成为电信诈骗的受害者.电信诈骗也不止是骗取钱财,犯罪分子还可以通过攻击大型公司获取数以亿计的用户信息,并用这些信息作为下一步的诈骗介质.

网络诈骗的低成本、低技术、易操作等特性使得网络诈骗层出不穷,同时网络诈骗方法的时效性使得诈骗方法极速的更新迭代,出现了各种形式的网络诈骗.本节着重介绍目前常见的诈骗形式,例如电话诈骗(telephone scam)[19-22]、问卷诈骗(online survey scam)[23]、移动广告诈骗(mobile advertisement scam)[24-26]、电子商务诈骗(e-commerce scam)[27]、技术支持诈骗(technical support scam)[20,28]、货物重运诈骗(reshipping scam)[29]、约会软件诈骗(online dating scam)[28,30]等,并对其所采用技术以及造成的危害进行归纳总结,对今后的研究工作进行展望.

1) 电话诈骗

电话作为人们日常使用最为普遍的通信渠道之一,自然成为了网络犯罪聚焦的重点.在打电话这一简单的过程中——主叫方拨出电话,经过电信运营商的路由到被叫方的SIM卡上接入,电信运营商和用户都可能成为电话诈骗的受害者.据通信诈骗管制协会(Communications Fraud Control Association, CFCA)估计,2019年对全球电信运营商由诈骗造成损失大约为283亿美元[31].

很多国家存在只能转售其他运营商服务的中小型运营商,其往往为了获取利益,会尝试劫持原本由其他运营商运营的电话.对此,Sahin等人[19-20]研究了国际收入分成诈骗(International Revenue Share Fraud, IRSF)的生态系统,并提出了对应的检测方法.

随着电话相关技术的发展,IP电话(VoIP)技术出现后,相关诈骗也逐渐增多.Sahin等人[22]分析了Over-The-Top(OTT)绕过这一新型诈骗形式.OTT指语音聊天软件(Skype等)利用IP网络通话服务即可连接到世界上任何地方[32].OTT绕过诈骗则是一种国际过境机构和OTT服务提供商未经主叫方、被叫方、运营商授权的情况下把正常拨出的电话通过IP转移至语音聊天软件的攻击.

IP电话的低成本的特点还催生了IP电话机器人的出现[33].传统诈骗电话的主叫方需要电话卡进行呼叫,费用较高,而IP电话机器人就可以自动化大规模低成本拨打诈骗电话.由于受害者通常在真的收到财产损失之后才会报告诈骗电话相关情况,导致大众无法即时了解最新的骗术.对此,Gupta等人[21]制作了大规模的电话蜜罐Phoneypot,对每个电话号码关联了年龄、地址等信息,吸引诈骗罪犯拨打这些号码,收集相关数据并进行分析.

2) 问卷诈骗

在线问卷是市场调研的重要步骤.通常,调查者只需在在线问卷网站上创建一份问卷,通过邮件或广告发送链接给用户,并搭配上金钱或实物奖励,就可以吸引用户进行填写.诈骗分子也注意到在线问卷市场的这一特点,通过丰厚奖励吸引用户填写问卷,事后却不兑现承诺的方式,进行欺诈活动.此外,部分攻击者还以邮寄奖品为由让用户填写敏感信息,或是用免费的礼品卡为诱饵让用户下载恶意软件、

勒索软件造成更大的损失.SURVEYLANCE[23]根据网站的内容、网络流量和页面的整体表现形式等信息构建分类器,在用户访问在线问卷时作为浏览器插件检查用户是否陷入了问卷诈骗,并通过大规模实验分析,检测出8 623个问卷诈骗网站.

3) 移动广告诈骗

手机应用丰富了人们的生活,应用内的移动广告则是免费应用的主要盈利方式.2020年,移动广告市场规模已扩大至1 870亿美元,占全球广告市场预算的30.5%[34].广告商根据广告呈现次数、用户点击量或应用安装量向广告服务提供商缴纳费用.移动广告诈骗的攻击者主要来自广告商的竞争对手,其主要目的是在用户没有被吸引的情况下耗尽广告商的推广费.

移动广告诈骗主要共有3种方法.对于按照用户点击量付费的广告商,一种方法是攻击者利用僵尸网络大量点击广告URL增加点击量的费用[35],二是攻击者欺骗用户点击不想看的广告[36]产生不必要的点击量.而对于按照广告呈现次数付费的广告商,攻击者则把广告隐藏在应用的其他元素下方,使得广告商在用户没有看到广告的情况下,依旧支付了广告费用.

MAdFraud[24]通过检测用户未交互时的HTTP请求来识别移动广告诈骗,并发现在130 339个安卓应用和35 087个恶意软件中有30%的应用会在后台运行时请求广告.Chen等人[25]关注了应用中通过webview实现的HTML广告,设计了一套监测移动广告生命周期的框架MADLIFE.Kim等人[26]基于AOSP[37]实现了一个动态测试移动广告诈骗的框架FraudDetective,利用堆栈跟踪从打开应用到提交广告诈骗活动之间的调用关系,检测是否有用户输入的存在.

4) 电子商务诈骗

电子商务是互联网发展后的一大应用,用户在电子商务平台购买产品时通常会根据产品的销量和评价决定是否购买商品.电子商务诈骗则是商家为了产生虚假的销量而逐渐形成的一套产业链[38].刷单诈骗是目前电子商务中广泛存在的一种诈骗形式:商家在刷单平台上招募工人,利用虚假的账户购买目标商品,但实际上没有真实的任何商品行为发生[39].另外,也有攻击者通过抢购促销商品,再加价转手获利[27].

电子商务诈骗依赖即时通信软件与工人交流,Wang等人[27]开发了Aubrey聊天机器人,把犯罪分子与工人的对话建模成有限状态机实现机器人在电子商务诈骗领域的自主对话,通过大量与犯罪分子对话来收集电子商务诈骗的信息.Aubrey通过加入了150个地下即时通信群聊,并与470名犯罪分子对话的方式,发现了38个售卖账号的地下市场和65个推广诈骗任务的附属网络.

5) 技术支持诈骗

当电脑出现问题时,人们习惯于寻找在线的技术支持来帮助解决问题,技术支持诈骗则针对这种场景产生.2019年美国联邦调查局网络犯罪投诉中心共收到48个国家13 633件来自受害者的技术支持诈骗投诉,损失总额超过5 400万美元,并较2018年增长了40%[40].在技术支持诈骗中,攻击者会创造一个带有流行软件和安全公司标志的网页,在用户访问时警告用户的设备已经感染了恶意软件,并诱导用户联系网页中显示的免费号码至“技术支持中心”寻求帮助.受害者通常会被要求下载远程桌面软件,允许远程技术人员通过软件操控设备.此时攻击者就会通过报错信息让受害者更加确信真的感染了恶意软件,要求受害者支付数百美元的恶意软件卸载费.除此之外,攻击者进一步可在受害者的设备上安装恶意软件窃取受害者的个人信息,进行后续攻击.

针对这一网络犯罪形式,Chen等人[28]基于网站的主机、网页大小和代码中链接数、关键词数等42个特征通过人工智能开发了技术支持诈骗检测系统AI@TTS.Miramirkhani等人[20]构建了一个分布式的爬虫(基础设施)来寻找涉及技术支持诈骗的域名,他们认为恶意广告是用户被吸引到技术支持诈骗页面的原因,通过从域名停放(domain parking)[41]和广告短链接爬取到的恶意广告访问可能存在的技术支持页面,并根据警示框、免费电话等特点找出其中的技术支持诈骗网站.

6) 货物重运诈骗

2013年底,Target(美国第二大折扣零售商)的数据遭到泄露,攻击者获得了多达1.1亿名客户的信用卡和个人信息[42].为了将被盗取的信用卡信息变现,犯罪分子使用被盗的信用卡从在线商店购买高价值的产品,然后将这些物品运送给以“在家工作”为幌子招募“马仔”.“马仔”再将收到的产品转发给远在国外的犯罪分子进行转售以获取非法利润[43].Hao等人[29]通过FBI的案件记录对这种诈骗产业链进行了分析.

7) 约会软件诈骗

人们习惯于通过约会软件认识新的朋友,然而

五花八门的约会软件中也出现了诈骗行为.Hu 等人[30]对诈骗约会软件进行了系统性的研究.诈骗约会软件通过广告吸引用户下载,注册步骤简单,且不需要任何个人信息,降低了用户的警惕性.在第 1 次登录后几分钟内就会有多名伪装成女性的机器人发来搭讪信息,并与受害者进行初步对话,如果想继续后续的对话就需要额外付费.付费后,就不会再有账户继续与受骗用户聊天.Hu 等人[30]通过关键词匹配和静态分析应用内付款 SDK 检测诈骗约会软件,在 9 个安卓应用市场的 250 万应用中检测出了 22 个家族的 967 个诈骗约会软件.

学术界对预防诈骗的方法进行了广泛讨论.对于诈骗而言,首要的预防方式是对用户的教育,许多诈骗方式(如诈骗约会软件、技术支持诈骗)对于用户而言没有复杂的技术,需要全社会的一起努力提高安全意识了解诈骗的危险性[23].对于通过网页进行诈骗的诈骗形式,浏览器可以提供帮助.如 Miramirkhani 等人[20]认为浏览器应能帮助用户离开如技术支持诈骗中的网页,提供一种关闭网页的快捷方式并在浏览器重启时不会重新打开.另外,大量的账号和电话号码等敏感信息是电子商务诈骗与电话诈骗的基础,遏制此类诈骗可以通过使用多重认证提高敏感信息注册和使用的门槛来打破犯罪产业链[27].对于货物重运诈骗,物流平台可以通过分析跟踪包裹信息的人员特点来及时制止还没运出的包裹[20].

目前对于诈骗的检测仍主要依赖于监督学习,攻击者可以通过规避监督学习所用的特征来逃避检测,未来可以发展半监督或无监督学习来检测约会软件诈骗[30]和问卷诈骗[23].同时,对于电子商务诈骗的信息收集目前使用的机器人是通过人为定义的有限状态机在与犯罪人员进行交流[27],在这一方向上未来可以通过深度学习等技术来探索更加通用的网络犯罪信息收集技术.另外,由于各大应用商店均不允许相关欺诈类软件进行上架传播,网络犯罪相关的应用软件分发方式也值得进一步研究.

1.3　恶意挖矿

恶意挖矿,是近来随着加密货币兴起而产生的一种全新的网络犯罪形式,即犯罪分子在未经许可的情况下侵占受害者的计算资源来挖取加密货币并获得利润.根据侵占计算资源方式的不同,可以将恶意挖矿分为针对浏览器的恶意挖矿与针对主机的恶意挖矿.其中由于针对浏览器的恶意挖矿影响范围更广、受害者更多,因此学术界对其的研究也更为深入.现代 Web 技术的发展使得浏览器具备了充分利用硬件资源的能力,利用这一特性,针对浏览器的恶意挖矿可以在未得到受害者的许可的情况下,通过网页内的 JavaScript 和 WebAssembly 与硬件资源直接进行交互进而隐蔽地占用计算资源挖掘加密货币.一般而言,针对浏览器的恶意挖矿主要有 4 种攻击方式:1)网站所有者主动发起恶意挖矿;2)犯罪分子在非法入侵网站后部署恶意挖矿脚本;3)通过入侵网站使用的第三方库、广告服务后分发恶意挖矿脚本;4)入侵路由器利用中间人攻击进行恶意挖矿.

为了检测恶意挖矿,研究人员往往从代码结构、网络请求、资源消耗等不同维度提取特征并进行检测,即在已知恶意挖矿常见的代码、行为特征的情况下,通过与待检测网站或脚本提取的特征进行匹配,如果发现目标的特征符合已知任意一类恶意挖矿的特征,则将其判定为恶意挖矿.基于对已知恶意挖矿的检测,研究人员往往会进一步对恶意挖矿的犯罪规模、非法收益等指标进行评估.本节将从恶意挖矿的检测方法,特别是检测时依赖的运行信息、程序行为、网络流量等特征,以及对恶意挖矿这种新型网络犯罪的规模评估这 2 方面介绍现有工作的研究成果.

在所有检测恶意挖矿的动态特征中,CPU 事件是最常用的特征,恶意挖矿往往会带来非常高的 CPU 占用.但由于大部分恶意挖矿会主动限制 CPU 占用,因此单独使用 CPU 占用作为特征必然会带来大量的漏报或误报,需要更多的其他特征来辅助更精准地检测.考虑到恶意挖矿不仅仅会进行大量的计算,还会频繁地读写内存、磁盘,与远端矿池进行交互,Ning 等人提出了工具 CapJack[44],通过提取 CPU 利用信息、内存信息、磁盘读写信息与网络流量信息作为特征,输入到分类器模型进行检测.此外,作为当代 CPU 的必要组成部分,在程序执行时硬件性能计数器会记录 CPU 内部执行的信息,如寄存器值、执行的指令等,这些信息同样能作为反映程序行为的特征.基于以上观察,Tahir 等人提出了基于硬件辅助分析的恶意挖矿检测方法[45],提取硬件性能计数器内的信息作为特征,然后使用随机森林作为分类器模型进行判别.

除了 CPU 信息的特征外,针对浏览器的恶意挖矿还会表现出特定的行为特征,比如周期性的哈希函数调用等.对此,Hong 等人提出了基于哈希函数的分析器和基于堆栈结构的分析器来记录恶意挖矿的行为表现[46].其中基于哈希函数的分析器会监控特定恶意挖矿常用哈希算法的出现频率,如门罗

币的 CryptoNight 算法,而基于堆栈结构的分析器则针对恶意挖矿中函数调用周期性的行为特征,校验相同函数调用出现的时间频率来分析调用的规律性.基于以上 2 类分析器提出的基于阈值的分析器经检验不存在任何误报.类似地,Kharraz 等人同样发现恶意挖矿中存在着特定的行为特征,比如脚本编译和执行的时间上存在着显著的差异,据此他们提出了工具 Outguard[47] 通过对浏览器进行插桩,Outguard 会收集网页内的资源加载时间、脚本编译时间、执行时间、垃圾回收信息,结合 CPU 的使用数据作为特征,然后使用支持向量机作为分类器模型进行检测.

为了更好地利用受害者机器的性能,针对浏览器的恶意挖矿往往会加载大量的 WebAssembly 文件.据此,Bian 等人提出了基于动态代码插桩的检测工具 MineThrottle[48],通过动态插入的性能计数器,MineThrottle 能精准分析不同代码块在运行时的 CPU 使用情况进而检测恶意挖矿现象.而 Rüth 等人从 WebAssembly 的文件特征入手,通过构建恶意挖矿常用 WebAssembly 的文件签名来检测浏览器中的恶意挖矿[49].无独有偶,Konoth 等人同样考虑到了 WebAssembly 文件在恶意挖矿中的重要作用,他们提出的检测工具 MineSweeper 通过计算 WebAssembly 中密码学相关指令的代码签名来识别特定的加密货币挖矿代码[50].此外,针对恶意挖矿的哈希算法利用 CPU 缓存的特性,MineSweeper 还会进一步监控 L1 和 L3 缓存来辅助判断.Naseem 等人则通过对 WebAssembly 进行分类的方式来检测恶意挖矿[51].他们提出的工具 MINOS* 将对恶意挖矿的检测问题转换为了对特定 WebAssembly 文件的二分类问题[12],针对任何加载了 WebAssembly 的网站,如果分类器确认了文件的恶意性,即可知对应网站存在恶意挖矿行为.

除了有效地检测恶意挖矿外,对恶意挖矿的影响规模、非法收益进行分析也是研究者关注的重点.Hong 等人通过网站平均访问量、停留时间、CPU 的平均哈希速率以及挖矿的平均收益估算得出,在 2018 年,恶意挖矿平均每月会影响超过 1 000 万以上的用户,平均每天额外消耗 27.8 万千瓦时的电能,为攻击者带来平均 5.9 万美元的日收入[46].Konoth 等人也用了类似的方法来估计恶意挖矿的收益,通过对网站访问时间和访问量的估计[50].除了估算恶意挖矿的收益外,Saad 等人还对比了恶意挖矿和在线广告的收益[52].他们的研究结果证明了在线广告的收益依旧远高于恶意挖矿的收益,因此从收益角度分析,绝大多数情况下恶意挖矿来自犯罪分子的非法入侵.

而 Bijmans 等人将目光放到了利用路由器进行中间人攻击的恶意挖矿上[53].在受害者使用被感染的路由器访问互联网时,犯罪分子会控制路由器在返回的内容中插入恶意挖矿的相关脚本,此时访问任意网站都会被攻击者劫持进行挖矿活动.此外,他们还通过先从互联网服务提供商获取区域内的全部流量信息,再从流量信息中提取全部可能被劫持的网络请求,结合每个请求返回网站的平均停留时间和 CPU 的平均哈希速率的方式估算该种攻击的收益情况.

为了分析恶意挖矿对互联网的整体影响,Bijmans 等人对互联网进行了 2 次不同规模采样和检测[54],发现高排名的网站中恶意挖矿现象更为普遍.此外,他们还发现检测出的恶意挖矿绝大多数与成人内容网站相关,而 2018 年 Hong 等人发现仅有 18%的恶意挖矿发生在成人内容相关的网站上,说明恶意挖矿行为随着时间的变化也在不断地扩散演进.

加密货币日益高涨的价格使得近年来恶意挖矿相关的网络犯罪不仅没有偃旗息鼓,反而更加猖獗,而且随着犯罪技术的不断演进,恶意挖矿的攻击方式和对抗手段也日趋复杂.可以看到从 2018 年初到目前,恶意挖矿的攻击手段已经从简单的在网站上部署挖矿脚本,通过内容混淆、限制使用率来对抗黑名单的检测,一步步发展为通过路由器漏洞的中间人攻击来分发恶意挖矿脚本的恶意挖矿,攻击原理更加复杂,且覆盖范围更广、非法所得更多.因此,针对恶意挖矿新的攻击形式和对抗方法的研究依旧会是未来恶意挖矿领域研究重点.

2 网络犯罪支撑技术

第 1 节中我们综述了网络犯罪产业链中网络犯罪攻击形式的犯罪逻辑和检测分析技术.在具体攻击形式的背后,通常有隐蔽的网络犯罪支撑技术对其提供保障与支持.网络犯罪支撑技术虽然不直接和用户产生接触,但其在为具体的攻击方式增加隐蔽性、提高攻击效率和成功率方面起着重要作用,如技术支持诈骗依赖于黑帽搜索引擎技术优化吸引用户访问;钓鱼网站依赖误植域名技术欺骗用户等.因此,近年来,对网络犯罪的支撑技术的研究也成为研究人员的关注重点.本节将以黑帽搜索引擎优化、误植

域名为例,论述网络犯罪支撑技术的概念、特点和检测分析方法.

2.1　黑帽搜索引擎优化

搜索引擎自诞生以来,极大地提升了用户从互联网上获取信息的有效性和便利性,据统计,每日人们会发起70亿余次的Google搜索请求[55].与此同时,搜索引擎提出了PageRank等为代表的算法对结果搜索结果进行不断优化.此外,随着需要被检索的信息增加,大部分搜索引擎也逐渐采用竞价排名来决定搜索结果的排名,热门关键词对应的价格水涨船高.此外,由于各地法律法规的制约,部分违法或灰色地带的关键词无法在搜索结果中呈现,使得非法产业只得另寻他法.高昂的价格加上法律法规的制约,促成了黑帽搜索引擎优化技术的产生和发展.

本节先简要介绍常见的黑帽搜索引擎优化技术,并讨论其对应的检测技术和评估方法,然后讨论当前黑帽搜索引擎检测技术上遇到的挑战并讨论后续发展方向.

黑帽搜索引擎优化指的是通过违规设置垃圾关键字、链接农场等方式,操纵针对特定关键词的搜索引擎结果的技术.目前广泛采用的技术包括:关键词篡改(keyword stuffing)、链接农场(link farm)、伪装(cloaking)等.

关键词篡改[56]即在网页的内容中填充一些与页面其余部分无关的流行词.攻击者希望通过将篡改的关键词与合法内容进行混合,提高网站在搜索引擎结果中的排行.为了最大限度地提高被索引的概率,通常一个网页中会嵌入数十甚至数百个无关的词.同时,搜索引擎也不断在更新自身的排名算法以应对关键词篡改的攻击,如Panda[57]和Penguin[58]都具有针对重复或篡改内容的页面的惩罚措施.然而,检测算法一般依赖于页面内容识别被操纵的网站.因此,黑帽攻击者也不断更新其相关技术,关键词篡改行为更加隐蔽以逃过检测.

链接农场[59-60]是指一组目的为了通过增加传入链接的数量来提高另一个网站的链接流行度的网站.其中,比较典型的链接农场类型包括私人博客[61-62]和论坛[63]等.整体而言,攻击者根据搜索引擎的算法规律,操纵权威网站组成的网络,通过建立反向链接的方式,将搜索权重传递给需要被提升排名的网站.

伪装[64-66]是指为了逃避检测,各类黑帽搜索引擎优化网站通常会向不同的访问者或搜索引擎动态提供不同的内容.例如通过浏览器的用户代理(user agent)、请求头(request header)、IP地址(IP address)等特征,保证只向目标客户呈现对应非法内容.

目前,学术界对于黑帽搜索引擎优化的研究主要集中于检测技术研究和威胁评估研究两大方面.

近年间,研究人员主要从搜索结果、网站内容、网页地图等角度研究、检测黑帽搜索引擎优化技术.文献[61,67-69]从搜索的结果作为入手点,其黑帽搜索引擎优化的最终目标是篡改搜索引擎结果.为此,不论攻击者采取何种技术,其结果一定可以在搜索引擎中发现.Liao等人[67]通过利用搜索结果和结果所在域名、搜索结果之间的不一致性,检测专属顶级域名中被被植入用于推广的网站.Joslin等人[68]利用通过语言学上的特点,系统性地生成并分析由于误拼写的关键词造成的问题.Wang等人[69]则通过搜索自动补全机制入手,采用词嵌入技术检测搜索引擎中关键词被黑帽搜索引擎优化污染的情况.Van等人[61]认为若网站需要实现搜索引擎优化的目标,则需要必须将自身结果进行收录.作者以此通过搜索引擎迭代反查种子集合,收集用于黑帽搜索引擎优化的私人博客网络样本.

与上述工作不同的,Yang等人[70]通过采用网站内容分析的角度进行研究,通过采用页面结构和网页内容相结合的方式进行恶意网站的识别检测;Du等人[71]则从网站地图(sitemap)入手,捕捉采用泛解析的黑帽搜索引擎优化网站.

近年间,为了对抗黑帽搜索引擎优化,各大搜索引擎也在频繁更新自身的排名策略.与之对应的,黑帽搜索引擎优化技术也普遍采用一些对抗方式.Yang等人[70]指出,目前搜索引擎可能利用人工容易识别、形状相近,但语义完全无关的的词语(如"六合彩"与"六台彩")躲避检测;Du等人[71]表明黑帽搜索引擎优化不但会采用动态内容生成来躲避搜索引擎的内容检测,还采用DNS的泛解析功能,动态生成域名以逃避搜索引擎的回环检测.文献[68,70]则表明部分黑帽SEO采用发音相近的字逃避检测.Yang等人[72]指出,攻击者可能采用"黑话"(jargon),逃避常见的关键词检测技术.

黑帽搜索引擎优化所带来的影响巨大,据Yang等人[70]报道,通过其对于7 000多个中文商业网站共超过3 800万个网页的持续观察,发现了在其中11%的网站被用于黑帽搜索引擎优化的行为;文献[62,73]则从时间维度上对黑帽SEO进行分析,Liao等人[74]识别出了3 186个云目录和318 470个门户页面用于常委关键词的黑帽SEO. Du等人[71]

通过对来自 22 个 TLD 和 SLD,共 1 350 万的域名的扫描,发现 458 个蜘蛛池网站,且虽然网站分布在超过 2.8 万个 IP 地址,但是其自治系统、域名注册商都集中分布在一小部分的 SEO 攻击者中.

从搜索结果的角度衡量,Joslin 等人[68]指出,谷歌和百度搜索引擎上,关于语言冲突搜索词的第一页搜索结果中,大约 1.19%指向恶意网站.文献[69]通过对 1.14 亿条搜索谷歌引擎的建议词进行分析,揭示了其中 0.48%的谷歌建议术语被操纵的现象,并且指出其中至少有 20%用于地下广告、宣传赌博内容甚至分发恶意软件.

作为一种支撑技术,黑帽搜索引擎优化的最终目的在于为对应的网络黑灰产引流,对于黑帽 SEO 所承载的攻击载荷,学术界也有相关工作对其进行研究.Liao 等人[67]揭露了针对.edu 域名的教育作弊攻击,此类网站主要提供教育相关的作弊服务,包括贩卖家庭作业,为学生提供在线考试作弊等服务.Joslin 等人[68]汇总了最受黑帽搜索引擎优化欢迎的前 5 个类别分别是:药品、成人相关、博彩、软件和汽车.Du 等人[71]则从 1 453 个已识别的黑帽搜索引擎优化客户站点中抓取的所有网页,并通过内容分析,发现使用黑帽搜索引擎优化的网站主要包括,销售和服务、赌博、代孕、新闻、色情、游戏、医疗和药品等.

虽然工业界和学术界普遍采用各种技术对抗黑帽搜索引擎优化攻击,但由于巨大的利益驱使,黑帽攻击者也在不断提高其自身的对抗手段.未来一段时间,检测技术和逃逸技术仍然会保持长期的螺旋上升态势.

对抗搜索引擎黑帽搜索引擎优化最简单的方法是配置有关域名、关键词的黑名单(blacklist).与之对应的,攻击者从最初的直接使用相关敏感词,逐渐转移到使用黑话,再到使用人工容易识别、但语义上完全无关的词语.为此,如何能在尽量少的基于专家人工经验的基础上准确生成关键词黑名单,尤其是先前未掌握的关键词黑名单,是未来研究的一个重点.

随着搜索引擎检测算法和非法页面检测工具的不断升级,传统的一次嵌入数百、数千个关键词的粗放式关键词篡改攻击也逐渐进化,出现了仅改变关键标签位置内容的攻击方式[70].同时,除了基于浏览器用户代理、请求头的伪装之外,还出现了基于动态代码执行的伪装技术等.因此,如何能在攻击日益隐蔽的情况下,持续准确识别黑帽搜索引擎优化的相关网站,也需要更多研究关注.

2.2 误植域名

误植域名(typosquatting)是一项历史悠久的网络犯罪技术,自从 20 世纪 90 年代域名开始注册以来,网络犯罪分子就开始通过抢先注册用户输入错误的域名,诱导其进入预设的网站产生访问流量.误植域名概念首先由 Gilwit 在《纽约法律杂志》[75]上提出,2003 年 Edelman[76]首次对误植域名进行了大规模的研究,发现超过 8 800 个误植域名,进一步调查发现大多数误植域名都属于 Zuccarini 所有.Zuccarini 经常针对儿童网站进行误植域名攻击,让误输域名的儿童重定向到色情网站从而牟利[76].

误植域名的检测首先要了解误植域名的生成模式.误植域名的生成模式经历了由最开始的简单字符变换到后续更高级的比特变换等模式的转变,相应的检测技术也在同步提升.研究人员们除了关注其中的检测技术以外,还关注这些误植域名网站背后的盈利模式和其他属性,以期望更加全面的了解这一灰色产业.本节也将从误植域名的生成模式、检测手段和其产业背后的其他相关属性进行梳理.

简单的误植域名生成模式可以分为 5 类[77-78].

① "."省略.输入域名忘记输入"."时发生,如由"www.example.com"变为"wwwexample.com".

② 字符省略.少输入域名字符时发生,如由"www.example.com"变为"www.exmple.com".

③ 字符错误排列.如由"www.example.com"变为"www.examlpe.com".

④ 字符替换.如由"www.example.com"变为"www.ezample.com".

⑤ 字符插入.如由"www.example.com"变为"www.exaample.com".

后续误植域名持续发展,出现了更加高级的生成模式:

⑥ 同态误植域名.Holgers 等人[79]提出了同态误植域名,即将目标域名变化为视觉上相似的域名.如"www.bankofthewest.com"变为"www.bankofthevvest.com",使用 2 个 v 来替换"w".

⑦ 比特误植域名(bitsquatting).Dinaburg 在黑帽安全会议上介绍了比特误植域名技术[80],这种技术不是依赖于用户的错误输入,而是由设备物理故障而发生的随机比特翻转错误.

⑧ 同音误植域名(soundsquatting).Nikiforakis 等人发现了同音误植域名[81],即利用单词的读音相似性来进行攻击,如"www.eight.com"变为"www.ate.com".

⑨ 组合误植域名(combosquatting).2017 年 Kintis 等人[82]研究了组合误植域名入技术,该技术将流行域名与单词连接起来,如"youtube.com"变为

“youtube-live.com”.Zeng 等人也将这类技术归为域名组合抢注[83].

误植域名技术的快速迭代也促使相应的检测技术不断发展.如表 1 所示,Wang 等人[77]设计了 Strider Typo-patrol 系统自动扫描并检测简单的误植域名(①～⑤),首先利用这些生成模式生成大量误植域名,再对误植域名进行访问判断是否被注册.Wang 等人这一检测方法是最早且被人们广泛引用的方法之一,如文献[84-85]中以同样的方法对误植域名的存活数量进行了检测.Holgers 等人[79]提出了同态误植域名的检测方法,即通过易被混淆的字符替换目标域名以生成误植域名,并利用 DNS 解析以确定域名是否被注册.Nikiforakis 等人[86]首次对比特误植域名现象进行了大规模的分析,提出了对应的检测方法.此后,Nikiforakis 等人[81]又提出了同音误植域名的检测方法:作者首先利用单词列表从域名中解析单词,后利用同音单词数据库来对目标域名单词进行变化生成误植域名.此后,Kintis 等人[82]第 1 次对组合误植域名进行了大规模研究,在长达 6 年的时间里,他们通过分析超过 4.68 亿条 DNS 记录识别出了 270 万个组合误植域名.上述误植域名的检测方法大都是以误植域名的生成模式为基础,自动化地生成大量的候选误植域名,最后利用 DNS 解析判断误植域名存在.根据他们实验结果表明,误植域名技术出现 20 多年来并没有随着时间演进而消亡,反而随着新技术的出现更加流行.

Table 1　Overview of Typosquatting-Related Research

表 1　误植域名相关研究概览

文献	目标域名数量	生成模式	待测误植域名数量	检出数量	检出率/%	发现
文献[77]	Alexa Top 10 000	①	10 000	5 094	51	
	Alexa Top 30	①～⑤	3 136	2 233	71	
	MillerSmiles Top 30	①～⑤	3 780	1 596	42	
	Top Children's Sites	①～⑤	7 094	2 685	38	
文献[79]		⑥		399		流行的网站具有更多已注册的同态误植域名;同态误植域名网站中诈骗网站数量最多.
文献[78]	Top 900*	①～⑤	约 3 000 000	1 050 000	35	误植域名的假网页比目标网页文件体积更小;流行网站的域名更受误植域名攻击的青睐.
文献[84]	Alexa Top 3 264	①～⑤	1 910 738	938 000	49	80%的误植域名网站通过点击付费或者竞争对手广告盈利;20%的误植域名通过重定向到其他站点盈利.
文献[86]	Alexa Top 500	⑦		5 366		Top500 网站受到比特误植域名攻击数量大致相同.
文献[85]	Alexa Top 1 000 000	①～⑤	约 4 700 000	940 000	20	95%的误植域名攻击的目标在于网站排名“长尾分布”中的处于“长尾”的网站.
文献[79]	Alexa Top 1 000	⑧	8 476	5 704	67	同音误植域名中有 56.88%的域名都是恶意的,其余多数属于保护性注册.
文献[89]	Alexa Top 500	①～⑤	28 179	17 172	16	误植域名攻击开始关注非流行网站;误植域名的网页在积极的切换其盈利方式.
文献[82]	Alexa 268	⑨		2 700 000		大多数组合误植域名不会长时间存在最大可达 1 000 天;组合误植域名相较于原始的误植域名更加普遍.

注:* 表示其作者并未明确指出使用的流行网站排名来源.

此外,在其他使用域名的场景,如邮件系统也会受到误植域名的影响,Szurdi 等人[87]首次对邮件的误植域名进行了研究,他们对自己注册的误植域名研究发现,这些误植域名对应的邮箱确实能够收到一些包含敏感信息的邮件,并且他们对真实世界中的邮件系统研究发现 1 211 个误植域名对应的邮箱每年会收到 800 000 个受害人的邮件.HTML 或者 JavaScript 代码中引入 JavaScript 库时也会出现由于误植域名而导致的安全问题,Nikiforakis[88]通过实验对这一攻击进行了研究,他们注册了与流行的 JavaScript 库名相近的域名,发现在 15 天之内有 16 万名开发者访问这些误植域名的库,这一攻击比传统的误植域名的威胁要大得多,可能引起恶意代码注入等安全问题.

除了对误植域名进行检测之外,研究人员们也关心误植域名的幕后经济模式以及其他相关属性.

2008 年 Baberjee 等人[78]对误植域名技术进行了大量的研究,他们发现误植域名的假网页比目标网页体积更小,到达假网页需要经过很多重定向耗费的时间更多,他们的研究结果也表明流行网站的域名更受误植域名攻击的青睐.7 年后,Szurdi 等人[85]得到了相反的结论,他们的研究表明 95%的误植域名攻击的目标在于网站排名"长尾分布"中的处于"长尾"的网站,Agten 等人[89]也证实了这一点,这表明误植域名的攻击趋势和对应行为随着时间在逐步变化.Agten 等人在长达 7 个月的研究中还发现误植域名的网页在积极的切换其盈利模式,有时利用广告有时通过诈骗犯罪进行盈利.文献[85,89]也证实目前误植域名仍然比较流行.近年来,Khan 等人[90]通过"意图推理"来量化误植域名技术对用户体验的影响.Spaulding 等人[91]通过用户调研的方式研究了误植域名技术在欺骗用户方面的有效性,研究发现误植域名攻击对熟悉安全相关知识的人效果不大,且字符替换和字符省略的误植域名相较其他误植域名攻击方式更有效.而 Tahir 等人[92]通过对人体手部结构、键盘布局和键字错误频率的探索发现,手部结构和键盘布局造成了某些字符组合拼写时更易出错,进而解释了导致域名更容易存在误植域名的现象.

误植域名技术更多的是用于为其他网络犯罪活动如钓鱼攻击[6]、网络诈骗[20]等提供技术支持.犯罪分子布置与目标网站相同的网页,在用户误植域名进入犯罪分子网站后进行钓鱼攻击.由于域名是用户主动输入的,用户在输入账户密码或者银行卡号等敏感信息时警惕心很低,很容易上当受骗.误植域名作为灰色产业地带,受到的相关法律法规监管较弱,所以首先需要从法律法规角度对这一块监管内容做出进步,同时由于用户粗心误植入错误的域名是不可避免的,对用户出现误植域名时进行提醒是有效保护用户利益的手段,如何做到用户误植域名的判断并及时提醒是今后研究的重点.

3 网络犯罪基础设施

仅有网络犯罪的攻击技术和支撑技术还不足以构成完整的网络犯罪产业链,攻击技术和支撑技术都需要对应的基础设施提供配套服务.例如,攻击者实施网络犯罪之前通过地下黑市进行相关信息或者攻击软件的获取,实施攻击钓鱼、诈骗等攻击时可以利用僵尸网络服务来进行高效便捷的内容分发,在攻击成功后利用洗钱渠道等手段来进行非法资金的变现等.这些网络犯罪基础设施无疑为网络犯罪攻击的成功实施和网络犯罪技术的有效支持提供了方便快捷的渠道.匿名、稳定、抗打击、能力强的基础设施是网络犯罪产业蓬勃发展的基石.本节对以地下论坛、僵尸网络和洗钱渠道为代表的网络犯罪产业链基础设施的相关概念、检测分析方法进行梳理.

3.1 地下论坛

地下论坛(underground forum)是网络犯罪分子买卖各种违法商品、服务或者信息的交易市场,一般也叫做地下黑市.地下黑市交易的违法商品如网络犯罪软件、个人身份信息、银行卡信息等,这些网络犯罪软件的大量交易使得实施网络犯罪的"门槛"逐渐降低,不具备安全背景的用户也可以通过"傻瓜式"操作来完成网络攻击,从而导致网络攻击事件频发,同时大量用户个人信息的交易也使得攻击的范围和规模逐步增大,威胁到大多数人的财产安全.

地下黑市在过去的 20 年间以多种形式存在.早期的地下黑市主要是利用 IRC 协议(Internet relay chat)通过群体聊天来交易[93-94],随着时间的发展网络犯罪分子逐渐将这些违法交易的场所转移到更加稳定、便利且流量更大的 Web 论坛[95-96].由于监管和打击的趋严,这些论坛逐步转入"地下",加入这些论坛有严格的审查过程,有些论坛至少需要 3 名现有成员推荐才能加入[97].并且某些地下论坛也逐步趋于专门化,如 Stone-Cross 等人[97]调查的 Spamdot.biz 论坛就专门售卖垃圾邮件服务.在地下黑市的研究工作中,研究人员们关注地下论坛中交换的非法信息和运作模式以估计其背后产生的经济效应.地下黑市类别众多变化极快,人工对这些数据分析成为了限制因素,所以研究人员们也在尝试探索自动化分析地下黑市的方法.本节从地下黑市的发展过程出发,梳理人们对地下黑市经济模式等属性和自动化分析方法的相关研究工作.

文献[93-94]中对早期基于 IRC 协议的地下黑市进行了研究,文献[93]中发现美国大约有 34~40 个的比较活跃的地下黑市,这些地下黑市一般都是基于 IRC 协议,他们通过多人聊天来完成商品的交易,基于 IRC 协议的论坛一般都是开放的很容易被发现,从而提醒监管者对这类地下黑市实施打击.Franklin 等人[94]首次大规模研究了基于 IRC 协议的地下黑市中商品的价格和类型以了解背后的经济模式,他们从活跃的地下论坛中收集了 7 个月的 1 300 万条聊天记录进行分析,发现这些地下市场中

热门的商品包括信用卡信息、金融信息和个人身份信息等.

相比基于 IRC 协议的地下黑市,基于 Web 论坛的地下黑市中的信息更加多样复杂,相应的每个地下黑市中商品的种类更多带来的经济规模也更加庞大.Zhuge 等人[95]首次对中国公共论坛形式的黑市进行了分析,他们提出了一种模型来描述游戏中虚拟资产等商品在的地下论坛中的具体交易流程,并从技术角度剖析了地下论坛影响下恶意网站构建过程的运作模式.Motoyama 等人[98]通过对 6 个地下论坛的数据进行分析,研究了地下市场社交网络的构成.Afroz 等人[99]利用在渔业和林业中大放异彩 OSTROM 经济框架探究了这些地下论坛的发展是否可持续,他们通过对 5 个地下论坛数据进行分析,发现可持续性取决于论坛的管理.Christin 等人通过在 2011—2012 年收集的论坛数据,对丝绸之路(Silk Road)这一典型的地下论坛进行了全面的研究分析[100].Pastrana[101]通过收集的 4 个地下论坛的长达 10 年的帖子,首次分析了地下市场中 10 年间交易货币的演变过程,发现亚马逊礼品卡也逐步成为地下黑市的交换媒介,且比特币是黑市中最受欢迎的交易货币.Hughes 等人[102]使用统计建模的方法分析了地下市场在 3 种不同的时期(动荡、平稳和新冠肺炎)中经济发展和社会特征,在动荡时期只有极少交易在增长;在平稳的时期他们观察到大范围的经济活动变化如亚马逊礼品卡这种中间货币大规模的兑换;在新冠肺炎大流行时期,他们发现所有类别的商品交易都在显著地增长.

地下黑市逐渐产生了专门化的趋势,研究人员们也开始对这些地下黑市中特定商品进行分析.Gross 等人[97]对地下论坛中垃圾邮件的经济模式进行分析,他们在主售垃圾邮件服务的 Spamdot. biz 论坛关闭之前收集了其中的数据,对其进行了全面的分析,研究发现该论坛中 Email 地址数据是最热销的商品,Sood 等人[103]揭示了地下市场中犯罪软件的买卖运作情况.Sun 等人[104]首次研究了地下市场中"特权滥用"现象,文献中的"特权滥用"类比于亚马逊等电商的买家通过商家设置商品的漏洞或者欺骗商家退款进行牟利.

为了减少研究地下黑市耗时费力的人力开销,研究人员们逐步转向探索自动化分析黑市的方法.Afroz 等人[105]首次探究了地下论坛中非结构数据的分析方法,他们利用 Stylometry 方法——通过分析写作风格来识别匿名信息的匿名作者,自动化识别地下论坛中用户的多重身份.Li 等人[106]则基于深度学习模型情绪分析方法分析其中顾客对卖家商品的反馈,从而自动化识别地下黑市的恶意软件或者病毒的制作者.Portnoff 等人[107]则利用自然语言处理技术(NLP)和机器学习实现了对地下黑市中帖子类别、商品种类和价格的识别的自动化识别.

地下黑市作为违法商品交易的场所,不仅为网络犯罪分子实施攻击提供了相应的信息和技术支持,也为他们犯罪后快速变现提供了方便的渠道.如地下黑市中售卖的垃圾邮件服务其中的受害者邮件地址或者分发邮件的僵尸网络都为垃圾邮件分发提供了更加高效方便的技术支持[97].如今大多数网络犯罪的实施者都不太精通安全技术,他们的网络犯罪攻击和攻击流程等信息都是从地下黑市中购买,然后加以简单操作便能实施高效的攻击,攻击之后再将所得个人信息等数据放到黑市中售卖.

网络黑市作为大多数网络攻击的源头和终点,对其整个生态或者产业链进行研究能够加深对网络犯罪动机等的理解,从源头上减少网络犯罪事件的发生,目前大多数研究都是做的这方面的工作.但是大多数对地下黑市的研究仍然停留在人工分析数据的阶段,探究地下黑市自动化分析方法的研究还是不足,所以探究如何自动化分析提取地下黑市的有价值信息,如何利用机器学习或者 NLP 技术抽象提取这些信息中更高维度的特征用于研究是后续的一个研究方向.同时目前对地下论坛的研究由于数据收集的限制侧重于比较开放的地下黑市,而对其他比较隐蔽甚至需要特殊协议进入的地下黑市研究甚少,对这部分黑市背后的产业链以及幕后经济模式的研究能够进一步加深对黑灰产行业的理解,也是未来需要研究的一个方向.

3.2　僵尸网络

僵尸网络(botnet)是在黑客命令控制下的一组受感染的终端主机,这些主机也被叫做"肉鸡".僵尸网络主要由 3 部分组成:僵尸网络控制者(botmaster)、命令控制通道(command and control)和僵尸主机(bots)[108].僵尸网络不仅本身能够造成巨大的危害,如进行 DDOS 攻击,还能为其他网络犯罪提供最基础的服务,如利用僵尸网络以垃圾邮件为载体进行钓鱼信息、诈骗信息和恶意软件等的传播.僵尸网络所涉及的研究范围较为广泛,本节将以垃圾邮件分发角度为例,梳理僵尸网络在网络犯罪生态系统中相关的技术研究.

在垃圾邮件的分发方面,垃圾邮件IP黑名单过滤机制很好地限制了垃圾邮件的传播,因此垃圾邮件发送者开始利用或者租用僵尸网络来进行垃圾邮件的分发[109].使用大量的僵尸主机进行垃圾邮件的传播显著提高了垃圾邮件投放的成功率和黑名单即时更新的难度[110].僵尸网络作为一个庞大的分布式计算网络,拥有很多僵尸主机,利用这些僵尸主机能够在数小时之内发送千万封垃圾邮件[111];在僵尸网络中,僵尸主机协同合作使得发送垃圾邮件的主机IP不断改变同时僵尸主机地理位置的多样性也使得它们很容易就逃避了垃圾邮件过滤技术和IP黑名单技术的检测.由于僵尸网络规模的庞大性和分发垃圾邮件的简便有效性,从2003年第1次使用僵尸网络分发垃圾邮件以来,僵尸网络已经成为发送垃圾邮件的主要方式,在2011年时使用僵尸网络来发放垃圾邮件的比例就已高达83.1%[112].

比较受欢迎的垃圾邮件僵尸网络服务有Bobax,Rustock,Storm等[113].Bobax是第一个基于模板的垃圾邮件僵尸网络,它每天发送90亿条垃圾邮件,有18.5万个发送垃圾邮件的僵尸主机[114].Rustock僵尸网络服务最早出现在2006年,它有15万到240万台僵尸主机,每小时能够发送多达2.5万封垃圾邮件[115],这些垃圾邮件涉及假药广告、假冒微软彩券诈骗等.Storm是最著名最大的僵尸网络,它在2007年利用100万~5 000万台受感染的主机发送了全球20%的垃圾邮件,这些垃圾邮件涉及到钓鱼攻击、雇佣诈骗广告等[116].

针对垃圾邮件僵尸网络服务的检测技术主要分为2种,基于被动发现的检测技术和基于主动发现的检测技术.基于被动发现的检测技术是指静默的观察僵尸网络发送垃圾邮件的活动并对产生的大量垃圾邮件进行深入分析,被动检测技术可以分为基于签名、基于DNS和基于异常的检测方式.基于签名的检测技术利用已知僵尸网络发布的垃圾邮件、恶意软件的签名或者指纹来进行检测,Xie等人[117]中提出了AutoRe框架,它以垃圾邮件的内容、服务器流量等属性来生成对应的指纹和特征,AutoRe对来自Hotmail的邮件进行随机抽样识别,发现了7 721起由僵尸网络发起的垃圾邮件活动.Ching等人[118]提出EIGENBOT,该技术可以通过基于语义图分析方法动态区分僵尸网络发送的垃圾邮件.基于DNS的检测技术通过监视DNS活动和检测不寻常的DNS查询来检测僵尸网络DNS流量.Ramachandran等人[119]中提出检查DNS黑名单(DNSBL)中域名的查询情况来寻找僵尸主机.其利用启发式规则来区分合法的DNSBL流量和来自僵尸网络的查询流量,然后利用僵尸网路的查询流量追踪僵尸主机.Ehrlich等人[120]提出了一种利用网络流数据和DNS元数据检测垃圾邮件源主机、僵尸主机以及僵尸网络控制者的方法.基于异常的僵尸网络检测是指,观察和分析不符合预期正常行为的电子邮件流量模式,以检测垃圾邮件的僵尸网络.Sroufe等人[121]提出了一种基于电子邮件“骨架”异常来分类僵尸网络的方法,“骨架”是指电子邮件HTML代码中每个标签中字符长度,然后利用机器学习将这种”骨架“信息用于分类不同僵尸网络发送的垃圾邮件.这种方式的一个局限性是,它不是检测垃圾邮件僵尸网络的完整解决方案,但它可以与垃圾邮件僵尸网络的网络行为分析相结合,以提高检测和分类僵尸网络的整体性能.Schafer等人[122]使用从SMTP服务器的日志文件中提取的数据来进行异常识别,他们从中提取源IP和连接时间来检测异常,从而检测被僵尸网络滥用的帐户.

基于主动发现的检测技术是通过创建一个模拟C&C协议的客户端从而加入到僵尸网络中,在加入僵尸网络后准确估计僵尸网络的大小,甚至破坏整个僵尸网络.进一步地,主动发现检测技术也可以分为3种类型,基于渗透的、基于受控环境的和基于Web恶意流量重定向的检测技术.在基于渗透的检测技术方面,Kreibich等人[123]利用分布式渗透方法对Storm僵尸网络分发的垃圾邮件进行了初步分析.Gross等人[97]通过主动攻击的方式获得Cutwail僵尸网络中13台僵尸网络服务主机的控制权,以此他们对僵尸网络中垃圾邮件的分发操作进行演示,同时他们也研究了基于IP的黑名单过滤机制的有效性、黑名单列表的质量和僵尸主机的可靠性等问题.在基于受控环境方面,Andreas等人[124]开发了Botnet judo系统,该系统将在受控环境中运行僵尸主机发送的垃圾邮件处理成正则表达式签名形式,然后进行垃圾邮件的实时检测.作者利用一种模板推理算法来产生垃圾邮件的正则表达式签名,并且在受控的僵尸主机收到命令发送垃圾邮件时对签名进行实时更新.作者还评估了多个模板推理的使用情况,证明judo在一种模板方式和多种模板方式下都很有效.在基于恶意流量重定向的检测技术方面,Ramachnaran和Feamster等人[125]研究了垃圾邮件的网络级特征,并观察了垃圾邮件发送者的网络级行为.通过对垃圾邮件传播数据的追踪,他们发现并确定了其分析的垃圾邮件来自BoBax僵尸网络.

目前存在很多对僵尸网络的研究工作,研究人员们提出的对僵尸网络的不同研究方式都有其局限性,特别是对于基于 P2P 协议的僵尸网络检测,目前基于主动的检测方式只能获得僵尸网络的一部分信息,使用被动探测和主动检测方式结合来探究 P2P 僵尸网络是今后的研究方向[113].同时随着移动设备、云服务、物联网的发展,会出现基于各种各样不同设备的僵尸网络攻击,这也对检测方法提出了更高的要求.

3.3　洗钱渠道

洗钱是犯罪分子用于逃避税务或者隐藏非法资金流向的手段[126].大多数网络犯罪都是以牟利为目的,网络犯罪分子一般最后都会进行资金的变现操作.但是使用常规的变现手段很容易被执法部门追踪资金流向进而导致账户被冻结,因此,犯罪分子会使用各种的洗钱手段来使所得资金"合法化".

传统的"物理"洗钱方式有如现金走私、赌博洗钱、保险洗钱和黑市比索交易(Balck Market Peso Exchange)等各种各样的方式,而现如今每天网络上发生着数亿笔交易,互联网的匿名性和不断发展的技术促使网上洗钱业务变成流行.洗钱是网络犯罪生态中关键的一步,随着支付机制的变化洗钱的方式也在逐步变化.网上洗钱方式从传统的使用 Liberty Reserve(一家在线支付公司)[127]洗钱、钱骡(Money mule)洗钱逐步转变到使用网络游戏虚拟货币洗钱和利用 PayPal[128]、比特币等货币的小额洗钱方式[129].洗钱已经作为一项服务在网络犯罪产业链的重要组成部分在地下论坛中广泛存在,因此研究人员们往往从地下论坛的数据作为切入点,研究洗钱过程中货币交换和资金流向等问题.本节从网上洗钱方式的发展和其中货币交换的角度梳理研究人员对网络犯罪生态中洗钱手段的研究.

Liberty Reserve 是一家支持匿名收付款的在线支付公司,很多网络犯罪分子曾利用其匿名性进行洗钱[129].但是自 2013 年 5 月 Liberty Reserve 被关闭后,该方式逐步被比特币洗钱替代[101].

钱骡是指利用自己的账户将非法资金或者高价值货物进行转移的中间人,他们在转移成功后会得到一笔报酬,由于将非法钱财转移到网络犯罪高发的一些国家存在困难,因此犯罪市场对钱骡存在大量需求[130].钱骡一般都是被以"在家高薪工作"等诈骗邮件为噱头欺骗的个体,他们并不知道其正在从事违法活动.根据 Mikhaylov 等人[130]的发现,学生、吸毒者、流浪汉和老人是犯罪分子们理想的钱骡招募目标.Hao 等人[29]发现了一种新型网络诈骗——货物重运诈骗(reshipping scams),犯罪分子用盗取的信用卡购买高价值商品之后通过地下市场服务租用"骡子"帮忙接收和运送包裹从而转移非法财产.

使用游戏虚拟货币也成为现在常用的洗钱手段,如某些大型网络游戏中存在虚拟获取兑换等功能,犯罪分子会利用非法资金购买游戏虚拟货币然后再将这虚拟货币在游戏论坛中售卖[131].其中网上赌博等游戏也属于这一类[129].Irwin 等人[132]提出对网络游戏大规模资金交易进行限制来对此类洗钱手段进行限制.

在洗钱过程中非法货币交换和资金流向方面,2016 年,Alexander 等人[130]对 2 个俄罗斯语的地下论坛进行了分析,发现俄罗斯网络犯罪分子更喜欢使用 Webmoney 电子商务支付系统[133]和西联国际汇款公司(Western Union)[134]来兑现或者进行资金的转移.然而,Portnoff 等人[107]证明现在比特币和 PayPal 是非法交易的首选货币,同时他们也观察到资金从 PayPal、比特币或者其他支付机制转移到信用卡的需求.Pastrana 等人[135]的研究也证明了 PayPal、比特币分别为网络犯罪经济中最常用的 2 种交易方式,同时他们也发现 2015 年后亚马逊礼品卡作为交易方式逐渐开始流行.Huang 等人[136]对勒索软件犯罪中支付的勒索金流向做了研究,他们发现勒索软件犯罪分子们更喜欢使用 BTC-e(俄罗斯一家将比特币兑换为法定货币的交易所)[137]来兑换比特币.最近 Siu 等人[138]研究发现 PayPal 仍然是网络犯罪分子最喜欢交换的交易方式,同时由于比特币价格的剧烈波动,非法资金交易媒介有从比特币转移到 PayPal 的趋势.

在网络犯罪生态中洗钱手段变换多样,目前已有很多研究对洗钱的方式进行探究,但是正如文献[129]中所阐述,洗钱方式会随着支付机制的转变而转变,因此对于新型洗钱手段的研究是今后的研究方向.同时洗钱变现作为网络犯罪过程中的最后一环,对洗钱过程的自动识别和溯源是为受害者挽回损失的基础,也是今后研究的重点之一.

4　总　　结

随着人们的生产生活对于网络依赖的提高.网络犯罪产业链规模日趋庞大,其攻击形式、支撑技术、基础设施的协调配合更加复杂.犯罪分子通过变化多端的攻击形式,损害人民群众的生命财产安全.

网络犯罪的攻击形式除了充分考虑心理学、社会工程学和人机交互设计特点之外,往往还有相关网络犯罪支撑技术为其提供相关技术支持,保证攻击的有效性和高效性.此外,不论是攻击形式和支撑技术均无法独立完成完整的犯罪过程,其需要通过对应的网络设施进行部署,在地下交易平台上买卖用户数据并通过洗钱渠道将非法所得"洗白".

针对网络犯罪产业链的复杂特点,本文从钓鱼、诈骗、恶意挖矿等经典的网络犯罪攻击形式入手,介绍网络犯罪攻击形式的特点和逻辑,进而综述黑帽搜索引擎优化、误植域名等网络犯罪支撑技术,并讨论了网络犯罪依赖的地下论坛、僵尸网络和洗钱渠道等基础设施,梳理网络犯罪生态的组成部分.最后,列举了网络犯罪研究中的一些仍存在的挑战和未来研究方向.

参 考 文 献

[1] Gordon S, Ford R. On the definition and classification of cybercrime [J]. Journal of Computer Virology and Hacking Techniques, 2006, 2(1): 13-20

[2] The Economist Newspaper. A history of hacking [OL]. [2021-08-24]. https://expectexceptional. economist. com/a-history-of-hacking.html

[3] Russell A. 56 people charged in India-based call centre scam that may have targeted Canadians. [OL]. [2021-08-24]. https://globalnews.ca/news/3032408/56-people-charged-in-india-based-call-centre-scam-that-may-have-targeted-canadians/

[4] Simonov N, Klenkina O, Shikhanova E. Leadingissues in cybercrime: A comparison of russia and japan [C] //Proc of the 6th Int Conf on Social, Economic, and Academic Leadership (ICSEAL-6-2019). Dordrecht: Atlantis Press, 2020: 504-510.

[5] Buil-Gil D, Miró-Llinares F, Moneva A, et al. Cybercrime and shifts in opportunities during COVID-19: A preliminary analysis in the UK [J]. European Societies, 2021, 23(S1): S47-S59

[6] Lin Yun, Liu Ruofan, Divakaran D M, et al. Phishpedia: A Hybrid Deep Learning Based Approach to Visually Identify Phishing Webpages [C] //Proc of the 30th USENIX Security Symp. Berkeley, CA: USENIX Association, 2021: 3793-3810

[7] Abdelnabi S, Krombholz K, Fritz M. Visualphishnet: Zero-day phishing website detection by visual similarity [C] //Proc of 2020 ACM SIGSAC Conf on Computer and Communications Security. New York: ACM, 2020: 1681-1698

[8] Cui Qian, Jourdan G V, Bochmann G V, et al. Tracking phishing attacks over time [C] //Proc of the 26th Int Conf on World Wide Web. New York: ACM, 2017: 667-676

[9] Tian Ke, Jan S T K, Hu Hang, et al. Needle in a haystack: Tracking down elite phishing domains in the wild [C] //Proc of the Internet Measurement Conf 2018. New York: ACM, 2018: 429-442

[10] Yoon C, Kim K, Kim Y, et al. Doppelgängers on the dark Web: A large-scale assessment on phishing hidden Web services [C] //Proc of the World Wide Web Conf. New York: ACM, 2019: 2225-2235

[11] Oest A, Zhang Penghui, Wardman B, et al. Sunrise to sunset: Analyzing the end-to-end life cycle and effectiveness of phishing attacks at scale [C] //Proc of the 29th USENIX Security Symp. Berkeley, CA: USENIX Association, 2020: 361-377

[12] Zhang Penghui, Oest A, Cho H, et al. CrawlPhish: Large-scale Analysis of Client-side Cloaking Techniques in Phishing [C] //Proc of 2021 IEEE Symp on Security and Privacy. Piscataway, NJ: IEEE, 2021 [2021-08-26]. https://www.kapravelos.com/publications/crawlphish-sp21.pdf

[13] Hu Hang, Jan S T K, Wang Yang, et al. Assessing Browser-level Defense against IDN-based Phishing [C] //Proc of the 30th USENIX Security Symp. Berkeley, CA: USENIX Association, 2021: 3739-3756

[14] Oest A, Safaei Y, Doupé A, et al. Phishfarm: A scalable framework for measuring the effectiveness of evasion techniques against browser phishing blacklists [C] //Proc of 2019 IEEE Symp on Security and Privacy (SP). Piscataway, NJ: IEEE, 2019: 1344-1361

[15] Oest A, Safaei Y, Zhang Penghui, et al. PhishTime: Continuous longitudinal measurement of the effectiveness of anti-phishing blacklists [C] //Proc of the 29th USENIX Security Symp. Berkeley, CA: USENIX Association, 2020: 379-396

[16] Maroofi S, Korczyński M, Duda A. Are you human? Resilience of phishing detection to evasion techniques based on human verification [C] //Proc of the ACM Internet Measurement Conf. New York: ACM, 2020: 78-86

[17] Oest A, Safei Y, Doupé A, et al. Inside a phisher's mind: Understanding the anti-phishing ecosystem through phishing kit analysis [C] //Proc of 2018 APWG Symp on Electronic Crime Research (eCrime). Piscataway, NJ: IEEE, 2018: 1-12

[18] Han Xiao, Kheir N, Balzarotti D. Phisheye: Live monitoring of sandboxed phishing kits [C] //Proc of 2016 ACM SIGSAC Conf on Computer and Communications Security. New York: ACM, 2016: 1402-1413

[19] Sahin M, Francillon A. Understanding and detecting international revenue share fraud [C] //Proc of the Network and Distributed System Security Symp (NDSS 2021). Reston, VA: The Internet Society, 2021 [2021-08-27]. https://www.ndss-symposium.org/wp-content/uploads/2021-051-paper.pdf

[20] Miramirkhani N, Starov O, Nikiforakis N. Dial one for scam: A large-scale analysis of technical support scams [C] //Proc of the Network and Distributed System Security Symp (NDSS 2017). Reston, VA: The Internet Society, 2017 [2021-08-27]. https://www.ndss-symposium.org/wp-content/uploads/2017/09/ndss2017_03B-1_Miramirkhani_paper.pdf

[21] Gupta P, Srinivasan B, Balasubramaniyan V, et al. Phoneypot: Data-drivenunderstanding of telephony threats [C] //Proc of the Network and Distributed System Security Symp (NDSS 2015). Reston, VA: The Internet Society, 2015 [2021-08-27]. https://www.ndss-symposium.org/wp-content/uploads/2017/09/03_2_3.pdf

[22] Sahin M, Francillon A. Over-the-top bypass: Study of a recent telephony fraud [C] //Proc of 2016 ACM SIGSAC Conf on Computer and Communications Security. New York: ACM, 2016: 1106-1117

[23] Kharraz A, Robertson W, Kirda E. Surveylance: Automatically detecting online survey scams [C] //Proc of 2018 IEEE Symp on Security and Privacy (SP). Piscataway, NJ: IEEE, 2018: 70-86

[24] Crussell J, Stevens R, Chen H. Madfraud: Investigating ad fraud in android applications [C] //Proc of the 12th Annual Int Conf on Mobile Systems, Applications, and Services. New York: ACM, 2014: 123-134

[25] Chen Gong, Meng Wei, Copeland J. Revisiting mobile advertising threats with MAdLife [C] //Proc of the World Wide Web Conf. New York: ACM, 2019: 207-217

[26] Kim J, Park J, Son S. The Abuserinside apps: Finding the culprit committing mobile ad fraud [C] //Proc of the Network and Distributed System Security Symp (NDSS 2020). Reston, VA: The Internet Society, 2020: 1-16

[27] Peng Wang, Liao Xiaojing, Qin Yue, et al. Into thedeep Web: Understanding e-commerce fraud from autonomous chat with cybercriminals [C] //Proc of the Network and Distributed System Security Symp (NDSS 2020). Reston, VA: The Internet Society, 2020 [2021-08-27]. https://www.ndss-symposium.org/wp-content/uploads/2020/02/23071.pdf

[28] Chen Yuchen, Chen Jiannliang, Ma Yiwei. AI @ TSS-Intelligent technical support scam detection system [J]. Journal of Information Security and Applications, 2021, 61(1): 13-17

[29] Hao Shuang, Borgolte K, Nikiforakis N, et al. Drops for stuff: An analysis of reshipping mule scams [C] //Proc of the 22nd ACM SIGSAC Conf on Computer and Communications Security. New York: ACM, 2015: 1081-1092

[30] Hu YangYu, Wang HaoYu, Zhou Yajin, et al. Dating with scambots: Understanding the ecosystem of fraudulent dating applications [J]. IEEE Transactions on Dependable and Secure Computing, 2019, 18(3): 1033-1050

[31] Communications Fraud Control Association. Fraud Loss Survey 2019 [OL]. [2021-08-16]. https://cfca.org/wp-content/uploads/2021/02/CFCA-2019-Fraud-Loss-Survey.pdf

[32] Green W, Lancaster B, Sladek J. Over the top services [J]. The Pipeline, 2006, 4(7): 11-17

[33] Tu Huahong, Doupé A, Zhao Ziming, et al. Sok: Everyone hates robocalls: A survey of techniques against telephone spam [C] //Proc of 2016 IEEE Symp on Security and Privacy (SP). Piscataway, NJ: IEEE, 2016: 320-338

[34] Zenith. Mobile share of advertising market to exceed 30% in 2020 [OL]. [2021-08-16]. https://www.zenithmedia.com/insights/global-intelligence-issue-06-2018/mobile-share-of-advertising-market-to-exceed-30-in-2020/

[35] Daswani N, Stoppelman M. The anatomy of Clickbot.A [C] //Proc of the 1st Conf on 1st Workshop on Hot Topics in Understanding Botnets. Berkeley, CA: USENIX Association, 2007 [2021-08-27]. https://www.usenix.org/legacy/events/hotbots07/tech/full_papers/daswani/daswani.pdf

[36] Zhang MingXue, Meng Wei, Lee S, et al. All your clicks belong to me: investigating click interception on the Web [C] //Proc of the 28th USENIX Security Symp. Berkeley, CA: USENIX Association, 2019: 941-957

[37] Android. Android Open Source Project [CP]. [2021-08-16]. https://source.android.com/

[38] Dow Jones & Company, Inc. The dark art of alibaba sales fakery [OL]. [2021-08-16]. https://www.wsj.com/articles/BL-CJB-26089

[39] Xu Haitao, Liu Daiping, Wang Haining, et al. E-commerce reputation manipulation: The emergence of reputation-escalation-as-a-service [C] //Proc of the 24th Int Conf on World Wide Web. New York: ACM, 2015: 1296-1306

[40] Federal Bureau of Investigation. 2019 Internet Crime Report [OL]. [2021-08-16]. https://pdf.ic3.gov/2019_IC3Report.pdf

[41] Alrwais S, Kan Yuan, Alowaisheq E, et al. Understanding thedark side of domain parking [C] //Proc of the 23rd USENIX Security Symp. Berkeley, CA: USENIX Association, 2014: 207-222

[42] The New York Times. For target, the breach numbers grow [OL]. [2021-08-16]. http://www.nytimes.com/2014/01/11/business/target-breach-affected-70-million-customers.html

[43] WAFF News. Re-shipping scam can turn job seekers into unwitting criminals [OL]. [2021-08-16]. http://www.waff.com/story/25034260/re-shipping-scam-can-turn-job-seekers-into-unwitting-criminals

[44] Ning Rui, Wang Cong, Xin ChunSheng, et al. Capjack: Capture in-browser crypto-jacking by deep capsule network through behavioral analysis [C] //Proc of the IEEE Conf on Computer Communications. Piscataway, NJ: IEEE, 2019: 1873-1881

[45] Tahir R, Durrani S, Ahmed F, et al. The browsers strike back: Countering cryptojacking and parasitic miners on the Web [C] //IEEE Conf on Computer Communications. Piscataway, NJ: IEEE, 2019: 703-711

[46] Hong Geng, Yang Zheming, Yang Sen, et al. How you get shot in the back: A systematical study about cryptojacking in the real world [C] //Proc of 2018 ACM SIGSAC Conf on Computer and Communications Security. New York: ACM, 2018: 1701-1713

[47] Kharraz A, Ma Z, Murley P, et al. Outguard: Detecting in-browser covert cryptocurrency mining in the wild [C] //Proc of the World Wide Web Conf. New York: ACM, 2019: 840-852

[48] Bian WeiKang, Meng Wei, Zhang MingXue. Minethrottle: Defending against wasm in-browser cryptojacking [C] //Proc of the World Wide Web Conf. New York: ACM, 2020: 3112-3118

[49] Rüth J, Zimmermann T, Wolsing K, et al. Digging into browser-based crypto mining [C] //Proc of the Internet Measurement Conf 2018. New York: ACM, 2018: 70-76

[50] Konoth R K, Vineti E, Moonsamy V, et al. Minesweeper: An in-depth look into drive-by cryptocurrency mining and its defense [C] //Proc of 2018 ACM SIGSAC Conf on Computer and Communications Security. New York: ACM, 2018: 1714-1730

[51] Naseem F, Aris A, Babun L, et al. MINOS*: A lightweight real-time cryptojacking detection system [C] //Proc of the Network and Distributed System Security Symp (NDSS 2021). Reston, VA: The Internet Society, 2021: 21-25

[52] Saad M, Khormali A, Mohaisen A. Dine and dash: Static, dynamic, and economic analysis of in-browser cryptojacking [C] //Proc of 2019 APWG Symp on Electronic Crime Research (eCrime). Piscataway, NJ: IEEE, 2019: 117-128

[53] Bijmans H L J, Booij T M, Doerr C. Just the tip of the iceberg: Internet-scale exploitation of routers for cryptojacking [C] //Proc of 2019 ACM SIGSAC Conf on Computer and Communications Security. New York: ACM, 2019: 449-464

[54] Bijmans H L J, Booij T M, Doerr C. Inadvertently making cyber criminals rich: A comprehensive study of cryptojacking campaigns at internet scale [C] //Proc of the 28th USENIX Security Symp. Berkeley, CA: USENIX Association, 2019: 1627-1644

[55] InternetLiveStats.com. Google Search Statistics [OL]. [2021-08-24]. https://www.internetlivestats.com/google-search-statistics/

[56] Ntoulas A, Najork M, Manasse M, et al. Detecting spam Web pages through content analysis [C] //Proc of the 15th Int conf on World Wide Web. New York: ACM, 2006: 83-92

[57] Google. Finding more high-quality sites in search [OL]. [2021-08-27]. https://googleblog.blogspot.com/2011/02/finding-more-high-quality-sites-in.html

[58] Aders A. What You Need to Know About Google's Penguin Update [OL]. [2021-08-27]. https://www.inc.com/aaron-aders/what-you-need-to-know-about-googles-penguin-update.html

[59] Chung Y, Toyoda M, Kitsuregawa M. A study of link farm distribution and evolution using a time series of Web snapshots [C] //Proc of the 5th Int Workshop on Adversarial Information Retrieval on the Web. New York: ACM, 2009: 9-16

[60] Wu BaoNing, Davison B D. Identifying link farm spam pages [C] //SpecialInterest Tracks and Posters of the 14th Int Conf on World Wide Web. New York: ACM, 2005: 820-829

[61] Van Goethem T, Miramirkhani N, Joosen W, et al. Purchasedfame: Exploring the ecosystem of private blog networks [C] //Proc of 2019 ACM Asia Conf on Computer and Communications Security. New York: ACM, 2019: 366-378

[62] Leontiadis N, Moore T, Christin N. A nearly four-year longitudinal study of search-engine poisoning [C] //Proc of 2014 ACM SIGSAC Conf on Computer and Communications Security. New York: ACM, 2014: 930-941

[63] Niu Yuan, Chen H, Hsu F, et al. A Quantitative Study of Forum Spamming Using Context-based Analysis [C] //Proc of the Network and Distributed System Security Symp (NDSS 2007). Reston, VA: The Internet Society, 2007 [2021-08-27]. https://www.ndss-symposium.org/wp-content/uploads/2017/09/A-Quantitative-Study-of-Forum-Spamming-Using-Context-based-Analysis-Yuan-Niu.pdf

[64] Wang D Y, Savage S, Voelker G M. Cloak and dagger: dynamics of Web search cloaking [C] //Proc of the 18th ACM Conf on Computer and Communications Security. New York: ACM, 2011: 477-490

[65] Chellapilla K, Chickering D M. Improvingcloaking detection using search query popularity and monetizability [C] //Proc of the 29th Annual Int ACM SIGIR Conf on Research and Development in Information Retrieval, SIGIR 2006. Bethlehem, PA: Lehigh University, 2006: 17-23

[66] Wu Baoning, Davison B D. Detecting semantic cloaking on the Web [C] //Proc of the 15th Int Conf on World Wide Web. New York: ACM, 2006: 819-828

[67] Liao Xiaojing, Yuan Kan, Wang Xiao Feng, et al. Seeking nonsense, looking for trouble: Efficient promotional-infection detection through semantic inconsistency search [C] //Proc of 2016 IEEE Symp on Security and Privacy (SP). Piscataway, NJ: IEEE, 2016: 707-723

[68] Joslin M, Li Neng, Hao Shuang, et al. Measuring and analyzing search engine poisoning of linguistic collisions [C] //Proc of 2019 IEEE Symp on Security and Privacy (SP). Piscataway, NJ: IEEE, 2019: 1311-1325

[69] Wang Peng, Mi Xianghang, Liao Xiaojing, et al. Game ofmissuggestions: Semantic analysis of search-autocomplete manipulations [C] //Proc of the Network and Distributed System Security Symp (NDSS 2018). Reston, VA: The Internet Society, 2018 [2021-08-27]. https://www.ndss-symposium.org/wp-content/uploads/2018/02/ndss2018_07A-2_Staicu_paper.pdf

[70] Yang Ronghai, Wang Xianbo, Chi Chi, et al. Scalable detection of promotional website defacements in black hat SEO Campaigns [C] //Proc of the 30th USENIX Security Symp. Berkeley, CA: USENIX Association, 2021: 3703-3720

[71] Du Kun, Yang Hao, Li Zhou, et al. The ever-changing labyrinth: A large-scale analysis of wildcard DNS powered blackhat SEO [C] //Proc of the 25th USENIX Security Symp. Berkeley, CA: USENIX Association, 2016: 245-262

[72] Yang Hao, Ma Xiulin, Du Kun, et al. How to learn klingon without a dictionary: Detection and measurement of black keywords used by the underground economy [C] //Proc of 2017 IEEE Symp on Security and Privacy (SP). Piscataway, NJ: IEEE, 2017: 751-769

[73] Antonakakis M, April T, Bailey M, et al. Understanding the mirai botnet [C] //Proc of the 26th USENIX Security Symp. Berkeley, CA: USENIX Association, 2017: 1093-1110

[74] Liao Xiaojing, Liu Chang, McCoy D, et al. Characterizing long-tail SEO spam on cloud Web hosting services [C] //Proc of the 25th Int Conf on World Wide Web. New York: ACM, 2016: 321-332

[75] Gilwit D B. The latest cybersquatting trend: Typosquatters, their changing tactics, and How to prevent public deception and trademark infringement [J]. Washington University Journal of Law & Policy, 2003, 11(1): 267-294

[76] Edelman B. Large-scale registration of domains with typographical errors [OL]. [2021-08-27]. https://cyber.harvard.edu/archived_content/people/edelman/typo-domains/

[77] Wang Yimin, Beck D, Wang J, et al. Strider typo-patrol: Discovery and analysis of systematic typo-squatting [C] // Proc of the 2nd Workshop on Steps to Reducing Unwanted Traffic on the Internet. Berkeley, CA: USENIX Association, 2006: 31-36

[78] Banerjee A, Barman D, Faloutsos M, et al. Cyber-fraud is one typo away [C] //Proc of the 27th IEEE Communications Society Conf on Computer Communications. Piscataway, NJ: IEEE, 2008: 1939-1947

[79] Holgers T, Watson D E, Gribble S D. Cutting through theconfusion: A measurement study of homograph attacks [C] //Proc of the USENIX Annual Technical Conf. Berkeley, CA: USENIX Association, 2006: 261-266

[80] Dinaburg A. Bitsquatting: DNS hijacking without exploitation [OL]. [2021-08-27]. http://dinaburg.org/bitsquatting.html

[81] Nikiforakis N, Balduzzi M, Desmet L, et al. Soundsquatting: Uncovering the use of homophones in domain squatting [C] //Proc of the Int Conf on Information Security. Berlin: Springer, 2014: 291-308

[82] Kintis P, Miramirkhani N, Lever C, et al. Hiding in plain sight: A longitudinal study of combosquatting abuse [C] // Proc of 2017 ACM SIGSAC Conf on Computer and Communications Security. New York: ACM, 2017: 569-586

[83] Zeng Yuwei, Chen Xunxun, Zang Tianning, et al. Windingpath: Characterizing the malicious redirection in squatting domain names [C] //Proc of the 22nd Int Conf on PAM 2021. Berlin: Springer, 2021: 93-107

[84] Moore T, Edelman B. Measuring the perpetrators and funders of typosquatting [C] //Proc of the Int Conf on Financial Cryptography and Data Security. Berlin: Springer, 2010: 175-191

[85] Szurdi J, Kocso B, Cseh G, et al. The long "taile" of typosquatting domain names [C] //Proc of the 23rd USENIX Security Symp. Berkeley, CA: USENIX Association, 2014: 191-206

[86] Nikiforakis N, Van Acker S, Meert W, et al. Bitsquatting: Exploiting bit-flips for fun, or profit? [C] //Proc of the 22nd Int Conf on World Wide Web. New York: ACM, 2013: 989-998

[87] Szurdi J, Christin N. Email typosquatting [C] //Proc of 2017 Internet Measurement Conf. New York: ACM, 2017: 419-431

[88] Nikiforakis N, Invernizzi L, Kapravelos A, et al. You are what you include: Large-scale evaluation of remote javascript inclusions [C] //Proc of 2012 ACM Conf on Computer and Communications Security. New York: ACM, 2012: 736-747

[89] Agten P, Joosen W, Piessens F, et al. Seven months' worth of mistakes: A longitudinal study of typosquatting abuse [C] //Proc of the 22nd Network and Distributed System Security Symp (NDSS 2015). Reston, VA: The Internet Society, 2015 [2021-08-27]. https://www.ndss-symposium.org/wp-content/uploads/2017/09/01_3_1.pdf

[90] Khan M T, Huo Xiang, Li Zhou, et al. Every second counts: Quantifying the negative externalities of cybercrime via typosquatting [C] //Proc of 2015 IEEE Symp on Security and Privacy. Piscataway, NJ: IEEE, 2015: 135-150

[91] Spaulding J, Upadhyaya S, Mohaisen A. You've been tricked! A user study of the effectiveness of typosquatting techniques [C] //Proc of the IEEE 37th Int Conf on Distributed Computing Systems (ICDCS). Piscataway, NJ: IEEE, 2017: 2593-2596

[92] Tahir R, Raza A, Ahmad F, et al. It's all in the name: Why some URLs are more vulnerable to typosquatting [C] //Proc of the IEEE Conf on Computer Communications. Piscataway, NJ: IEEE, 2018: 2618-2626

[93] Cymru T. The underground economy: priceless [J]. login, 2006, 31(6): 7-16

[94] Franklin J, Perrig A, Paxson V, et al. An inquiry into the nature and causes of the wealth of internet miscreants [C] // Proc of the 14th ACM Conf on Computer and Communications Security. New York: ACM, 2007: 375-388

[95] Zhuge Jianwei, Holz T, Song ChengYu, et al. Studying malicious websites and the underground economy on the Chinese Web [M] //Managing Information Risk and the Economics of Security. Berlin: Springer, 2009: 225-244

[96] Radianti J. A study of a social behavior inside the online black markets [C] //Proc of the 4th Int Conf on Emerging Security Information, Systems and Technologies. Piscataway, NJ: IEEE, 2010: 189-194

[97] Stone-Gross B, Holz T, Stringhini G, et al. The underground economy of spam: a botmaster's perspective of coordinating large-scale spam campaigns [C] //Proc of the 4th USENIX Conf on Large-scale Exploits and Emergent Threats. Berkeley, CA: USENIX Association, 2011: 4-12

[98] Motoyama M, McCoy D, Levchenko K, et al. An analysis of underground forums [C] //Proc of 2011 ACM SIGCOMM Conf on Internet Measurement Conf. New York: ACM, 2011: 71-80

[99] Afroz S, Garg V, McCoy D, et al. Honor among thieves: A common's analysis of cybercrime economies [C] //Proc of 2013 APWG eCrime Researchers Summit. Piscataway, NJ: IEEE, 2013: 77-87

[100] Christin N. Traveling the Silk Road: A measurement analysis of a large anonymous online marketplace [C] //Proc of the 22nd Int Conf on World Wide Web. New York: ACM, 2013: 213-224

[101] Pastrana S, Thomas D R, Hutchings A, et al. Crimebb: Enabling cybercrime research on underground forums at scale [C] //Proc of 2018 World Wide Web Conf. New York: ACM, 2018: 1845-1854

[102] Vu A V, Hughes J, Pete I, et al. Turning up the dial: the evolution of a cybercrime market through set-up, stable, and covid-19 eras [C] //Proc of the ACM Internet Measurement Conf. New York: ACM, 2020: 551-566

[103] Sood A K, Enbody R J. Crimeware-as-a-service—A survey of commoditized crimeware in the underground market [J]. International Journal of Critical Infrastructure Protection, 2013, 6(1): 28-38

[104] Sun Zhibo, Oest A, Zhang Penghui, et al. Havingyour cake and eating it: An analysis of concession-abuse-as-a-service [C] //Proc of the 30th USENIX Security Symp. Berkeley, CA: USENIX Association, 2021: 4169-4186

[105] Afroz S, Islam A C, Stolerman A, et al. Doppelgänger finder: Taking stylometry to the underground [C] //Proc of 2014 IEEE Symp on Security and Privacy. Piscataway, NJ: IEEE, 2014: 212-226

[106] Li Weifeng, Chen H. Identifying top sellers in underground economy using deep learning-based sentiment analysis [C] // Proc of 2014 IEEE Joint Intelligence and Security Informatics Conf. Piscataway, NJ: IEEE, 2014: 64-67

[107] Portnoff R S, Afroz S, Durrett G, et al. Tools for automated analysis of cybercriminal markets [C] //Proc of the 26th Int Conf on World Wide Web. New York: ACM, 2017: 657-666

[108] Feily M, Shahrestani A, Ramadass S. A survey of botnet and botnet detection [C] //Proc of the 3rd Int Conf on Emerging Security Information, Systems and Technologies. Piscataway, NJ: IEEE, 2009: 268-273

[109] Pathak A, Qian F, Hu Y C, et al. Botnet spam campaigns can be long lasting: evidence, implications, and analysis [J]. ACM SIGMETRICS Performance Evaluation Review, 2009, 37(1): 13-24.

[110] Abu Rajab M, Zarfoss J, Monrose F, et al. A multifaceted approach to understanding the botnet phenomenon [C] // Proc of the 6th ACM SIGCOMM Conf on Internet Measurement. New York: ACM, 2006: 41-52

[111] Zhu Zhaosheng, Lu Guohan, Chen Yan, et al. Botnet research survey [C] //Proc of the 32nd Annual IEEE Int Computer Software and Applications Conf. Piscataway, NJ: IEEE, 2008: 967-972

[112] Blasczyk S, Piontek M. Symantec Intelligence Report: November 2011 [OL]. [2021-00-00]. https://docs.broadcom.com/doc/intelligence-report-nov11-en

[113] Khan W Z, Khan M K, Muhaya F T B, et al. A comprehensive study of email spam botnet detection [J]. IEEE Communications Surveys & Tutorials, 2015, 17(4): 2271-2295

[114] Wikipedia. Bagle [OL]. [2021-08-22]. http://en.wikipedia.org/wiki/Bagle

[115] Wikipedia. Rustock botnet [OL]. [2021-08-22]. http://en.wikipedia.org/wiki/Rustock_botnet

[116] Wikipedia. Storm botnet [OL]. [2021-08-22]. http://en.wikipedia.org/wiki/Storm_botnet

[117] Xie Yinglian, Yu Fang, Achan K, et al. Spamming botnets: signatures and characteristics [J]. ACM SIGCOMM Computer Communication Review, 2008, 38(4): 171-182

[118] Mao Chinghan, Lin Changchang, Pan J Y, et al. EigenBot: Foiling spamming botnets with matrix algebra [C] //Proc of the ACM SIGKDD Workshop on Intelligence and Security Informatics. New York: ACM, 2012: 35-42

[119] Ramachandran A, Feamster N, Dagon D. Revealing botnet membership using dnsbl counter-intelligence [C] //Proc of the 2nd Workshop on Steps to Reducing Unwanted Traffic on the Internet. Berkeley, CA: USENIX Association, 2006: 49-54

[120] Ehrlich W K, Karasaridis A, Hoeflin D A, et al. Detection ofspam hosts and spam bots using network flow traffic modeling [C] //Proc of the 3rd USENIX Workshop on Large-Scale Exploits and Emergent Threats: Botnets, Spyware, Worms, and More. Berkeley, CA: USENIX Association,2010: 7-15

[121] Sroufe P, Phithakkitnukoon S, Dantu R, et al. Email shape analysis for spam botnet detection [C] //Proc of the 6th IEEE Consumer Communications and Networking Conf. Piscataway, NJ:IEEE, 2009: 423-424

[122] Schäfer C. Detection of compromised email accounts used by a spam botnet with country counting and theoretical geographical travelling speed extracted from metadata [C] // Proc of 2014 IEEE Int Symp on Software Reliability Engineering Workshops. Piscataway, NJ: IEEE, 2014: 329-334

[123] Kreibich C, Kanich C, Levchenko K, et al. On the spam campaign trail [C] //Proc of the 1st USENIX Workshop on Large-Scale Exploits and Emergent Threats. Berkeley, CA: USENIX Association, 2008 [2021-09-13]. https://www.usenix.org/legacy/events/leet08/tech/full_papers/kreibich/kreibich.pdf

[124] Pitsillidis A, Levchenko K, Kreibich C, et al. Botnet Judo: Fighting Spam with Itself [C] //Proc of the Network and Distributed System Security Symp (NDSS 2010). Reston, VA: The Internet Society, 2010 [2021-08-27]. https://www.ndss-symposium.org/wp-content/uploads/2017/09/pits.pdf

[125] Ramachandran A, Feamster N. Understanding the network-level behavior of spammers [C] //Proc of 2006 Conf on Applications, Technologies, Architectures, and Protocols for Computer Communications. New York: ACM, 2006: 291-302

[126] Levi M, Reuter P. Money laundering [J]. Crime and Justice, 2006, 34(1): 289-375

[127] Wikipedia. Liberty Reserve [OL]. [2021-08-24]. https://en.wikipedia.org/wiki/Liberty_Reserve

[128] PayPal. PayPal [OL]. [2021-08-24]. https://www.paypal.com/

[129] Richet J L. Laundering Money Online: A review of cybercriminals methods [OL]. [2021-08-27]. https://arxiv.org/abs/1310.2368

[130] Mikhaylov A, Frank R. Cards, money and two hacking forums: An analysis of online money laundering schemes [C] //Proc of 2016 European Intelligence and Security Informatics Conf (EISIC). Piscataway, NJ: IEEE, 2016: 80-83

[131] Moiseienko A, Izenman K. Gaming thesystem: Money laundering through online games [J]. Rusi Centre for Financial Crime and Security Studies, AML and CTF, 2019, 39(9): 1-5

[132] Irwin A S M, Slay J, Choo K K R, et al. Money laundering and terrorism financing in virtual environments: A feasibility study [J]. Journal of Money Laundering Control, 2014, 17(1): 50-75

[133] WebMoney. WebMoney [OL]. [2021-08-24]. https://www.wmtransfer.com/

[134] Western Union Holdings, Inc. Western Union [OL]. [2021-08-24]. https://www.westernunion.com/cn/en/home.html

[135] Pastrana S, Hutchings A, Caines A, et al. Characterizing eve: Analysing cybercrime actors in a large underground forum [C] //Proc of Int Symp on Research in Attacks, Intrusions, and Defenses. Berlin: Springer, 2018: 207-227

[136] Huang D Y, Aliapoulios M M, Li V G, et al. Tracking ransomware end-to-end [C] //Proc of 2018 IEEE Symp on Security and Privacy (SP). Piscataway, NJ: IEEE, 2018: 618-631

[137] Wikipedia. BTC-e [OL]. [2021-08-24]. https://en.wikipedia.org/wiki/BTC-e

[138] Siu G A, Collier B, Hutchings A. Follow the money: The relationship between currency exchange and illicit behaviour in an underground forum [C] //Proc of the Workshop on Actors and Cyber-Crime Operations (IEEE European Symp on Security and Privacy 2021). Piscataway, NJ: IEEE, 2021 [2021-08-27]. https://www.cl.cam.ac.uk/~ah793/papers/2021follow_the_money.pdf

Hong Geng, born in 1994. PhD candidate. His main research interests include cybercrime and mobile security.

洪　赓,1994 年生.博士研究生.主要研究方向为网络犯罪、移动应用安全.

Yang Sen, born in 1997. Master candidate. His main research interests include cybercrime and mobile security.

杨　森,1997 年生.硕士研究生.主要研究方向为网络犯罪、移动应用安全.

Ye Han, born in 1998. Master candidate. His main research interests include mobile privacy and mobile security.

叶　瀚,1998 年生.硕士研究生.主要研究方向为移动应用隐私保护、移动应用安全.

Yang Zhemin, born in 1984. PhD, associate professor, master supervisor. His main research interests include operation system, system security, cybercrime, program analysis and runtime environment technology.

杨哲慜,1984 年生.博士,副教授,硕士生导师.主要研究方向为操作系统、系统安全、程序分析和运行环境技术.

Yang Min, born in 1979. PhD, professor, PhD supervisor. Senior member of CCF. His main research interests include network security, vulnerability exploiting and analysis, AI security, cybercrime and system security mechanism.

杨　珉,1979 年生.博士,教授,博士生导师,CCF 高级会员.主要研究方向为网络安全、漏洞挖掘和分析、人工智能安全、网络犯罪和系统安全机制.