

电子邮件安全扩展协议应用分析

尚菁菁^{1,2,3}, 朱宇佳^{2,3}, 刘庆云^{2,3}

(1. 中国科学院大学网络空间安全学院, 北京 100093;

2. 信息内容安全技术国家工程实验室, 北京 100093;

3. 中国科学院信息工程研究所, 北京 100093)

摘要: 电子邮件是黑客发起网络攻击的主要入口, 其中身份仿冒是电子邮件欺诈重要手段。基于邮件身份验证机制, 构建属性图以测量政府机构电子邮件安全扩展协议全球采用率。从邮件内容仿冒、域仿冒、信头仿冒 3 个维度研究安全扩展协议部署效果。结果表明, 各国政府机构邮件系统中部署 SPF 协议的约占 70%, 部署 DMARC 协议的不足 30%, 电子邮件身份检测采用率较低。当欺诈邮件进入收件人邮箱后, 邮件服务提供商针对仿冒邮件警告机制有待完善。

关键词: 电子邮件; 安全扩展协议; 域名密钥识别邮件标准; 发件人策略框架; 基于域的邮件验证、报告和一致性

中图分类号: TP393

文献标识码: A

doi: 10.11959/j.issn.2096-109x.2020083

Analysis of security extension protocol in e-mail system

SHANG Jingjing^{1,2,3}, ZHU Yujia^{2,3}, LIU Qingyun^{2,3}

1. School of Cyber Security, University of Chinese Academy of Sciences, Beijing 100093, China

2. National Engineering Laboratory for Information Security Technologies, Beijing 100093, China

3. Institute of Information Engineering, Chinese Academy of Sciences, Beijing 100093, China

Abstract: E-mail is the main entry point for hackers to launch network attacks. Impersonating a trusted entity is an important means of e-mail forged. An attribute graph based on the e-mail authentication mechanism was built to measure the global adoption rate of e-mail security extension protocols for government agencies. Research on the deployment effect of security extension protocol was from three dimensions: e-mail content phishing, domain phishing, and letterhead phishing. The results show that about 70% of the SPF protocols are deployed in the mail systems of government agencies in various countries, and less than 30% of the DMARC protocol is deployed. The adoption rate of email identity detection is low. When forged e-mail gets in, the e-mail providers' warning mechanism for counterfeit emails need to be improved.

Key words: e-mail, security extension protocol, DKIM, SPF, DMARC

收稿日期: 2020-05-06; 修回日期: 2020-07-03

通信作者: 朱宇佳, zhuyujia@iie.ac.cn

基金项目: 中国科学院战略性先导科技专项 (XDC02030000); 国家重点研发计划 (2016YFB0801300)

Foundation Items: Chinese Academy of Sciences Strategic Pilot Project (XDC02030000), National Key R & D Program (2016YFB0801300)

论文引用格式: 尚菁菁, 朱宇佳, 刘庆云. 电子邮件安全扩展协议应用分析[J]. 网络与信息安全学报, 2020, 6(6): 69-79.

SHANG J J, ZHU Y J, LIU Q Y. Analysis of security extension protocol in e-mail system[J]. Chinese Journal of Network and Information Security, 2020, 6(6): 69-79.

1 引言

邮件服务是互联网最基础的服务之一，在互联网通信中有着不可替代的作用。近年来，鱼叉式网络钓鱼持续威胁用户和组织，账号密码被钓鱼等恶性事件层出不穷。根据 2019 年赛门铁克《互联网威胁安全报告》，每 412 封电子邮件中就有 1 封包含恶意软件攻击，每月有 7 710 个组织受到“商务电子邮件”攻击。根据沃通报告，2019 年商务电子邮件泄露攻击给组织机构造成约 17.7 亿美元的损失，针对商务电子邮件攻击，黑客最常用的手段就是冒充企业高管给员工发送电子邮件，指示员工执行某项操作。2020 年 2 月 7 日，国家互联网应急中心发出预警，通报网络不法分子利用 2019 新型冠状病毒相关题材，冒充国家卫生健康委员会等疫情防控部门，向我国部分单位和用户投放与新型肺炎疫情相关的钓鱼邮件。造成这些恶性事件的原因大多数是攻击者应用欺骗技术模仿受信任的实体，以进行具有高度欺骗性的网络钓鱼攻击，因此电子邮件身份验证十分重要。

针对电子邮件身份验证主要有 3 个安全扩展协议：发件人策略框架（SPF，sender policy framework）、域名密钥识别邮件标准（DKIM，domain keys identified mail）和基于域的邮件验证、报告和一致性（DMARC，domain-based message authentication, reporting and conformance）。SPF 协议负责检查发件人是否仿冒了邮箱域名，DKIM 协议不仅可以检查发件人身份信息，还可以检测邮件内容是否被篡改，然而 SPF 和 DKIM 协议只负责检测发件人身份，对于检测结果的处理并不明确，狡猾的攻击者很快发现并利用这些安全协议的不足之处，继续制造和发送钓鱼邮件。DMARC 协议是基于现有的 SPF 和 DKIM 两大主流电子邮件安全协议发展而来的，与上面两种协议不同的是，DMARC 协议在检测发件人身份信息后要求生成一份检测结果报告发送到指定的管理员邮箱，以便管理员对钓鱼邮件及发件人信息做定期检查处理。基于 3 种安全扩展协议，可以确保消息的真实性，防止攻击者篡改或仿冒电子邮件。然而，2015 年进行的

Realworld 测量显示，在 Alexa Top1 Million 域中，仅有 40% 具有 SPF 协议，1% 具有 DMARC 协议，并且很多配置上存在格式错误或缺少版本信息等问题^[1-2]。

在政企类机构中，邮件服务作为主要的交互手段，其承载了大量的有价值信息，是黑客发起网络攻击和信息窃取的主要入口，因此对于政企类机构，验证传入电子邮件的真实身份信息尤为重要。为了分析特定企业邮箱域名安全扩展协议部署策略，需要全面获取邮箱域名并对邮箱域名属性进行标识。邮箱域名发现与标识方法有主动采集、被动获取、主被动协同 3 种。主动采集法通过网络爬虫直接从相关网站上获取所需数据，获取的数据可知可溯源，易于标识属性信息。但该方法多数情况下只能获得企业网站域名，发现的数据有限。被动获取的方法全面依赖于网络中已有的流量，涉及范围大，数量多，可以对主动采集中缺少的数据予以补充，通过特征匹配识别邮箱域名与网站域名之间的关系。但被动数据关联性弱，缺少邮箱域名属性信息。主被动协同的方法可以结合主动测量可知可溯源性及被动测量的广泛性全面采集并标识邮箱域名。

本文主要研究现有安全扩展协议安全配置现状。首先基于主被动协同的方法，多表征融合标识邮箱域名。其次，基于邮件信源认证机制，对政府机构所使用的电子邮件服务提供商的安全扩展协议的采用率进行测量分析比较。最后，从邮件内容仿冒、域仿冒、信头仿冒 3 方面研究构建电子邮件仿冒模型，模拟端到端的仿冒邮件实验，探究邮件服务提供商如何在电子邮件可用性与安全性之间进行权衡，及邮件服务提供商检测策略与客户端真实通知之间的差距，全面分析邮件安全扩展协议部署效果。

本文的主要贡献有 3 项。

- 1) 基于主被动协同的方法多表征融合标识邮箱域名。
- 2) 基于属性图标识邮件安全扩展协议部署策略。
- 3) 基于欺诈邮件原理，提出多维度仿冒邮件策略，揭示电子邮件提供商如何处理仿冒电子邮件，如何在电子邮件可用性和安全性之间达到权衡。

2 相关工作

电子邮件系统是基于简单邮件传输协议 (SMTP, simple mail transfer protocol)^[1]构建的, 该协议最初设计时并未考虑安全性。后来引入了众多安全扩展协议, 以提供机密性、完整性和真实性^[2]。本文主要关注电子邮件的真实性, 网络钓鱼攻击者通常滥用该属性对用户进行欺诈攻击。下面将首先介绍电子邮件真实性检测协议发展历程及本文测量涉及的电子邮件安全扩展协议, 然后介绍电子邮件安全扩展协议的研究现状及不足。

2.1 电子邮件安全协议发展历程

1982年, IETF首先提出SMTP, 用于用户之间纯文本消息传输^[1]。为提供多种类型消息传输, IETF于1996年颁布了多用途互联网邮件扩展 (MIME, multipurpose internet mail extensions)^[3], 使用户之间可发送和接收音视频、图片、附加文件等多种数据格式的邮件。为了增强电子邮件传输的机密性和完整性, IETF于1996年和1999年先后提出基于MIME的扩展协议优良保密协议 (PGP, pretty good privacy)^[4]和多用途网际邮件扩充协议 (S/MIME, secure multipurpose internet mail extensions)^[5]实现端到端的电子邮件加密传输。同时1999年提出了STARTTLS安全协议^[6], 将SMTP升级为加密传输。目前电子邮件加密协议已广泛受到各邮件服务提供商的关注。然而, 上述这些协议的提出只能防止电子邮件遭受攻击者被动监听, 攻击者仍可仿冒发件人身份进行网络钓鱼攻击。为了确保电子邮件的真实性, IETF于2000年和2014年分别起草并标准化了发件人策略框架SPF协议^[7], 用于查询发送域经过的SMTP服务器是否在发送域指定允许的IP段内。2004年和2011年又先后起草并确认了DKIM协议^[8], 不仅可以检查发件人身份信息, 还能检测发件内容是否被篡改。随后2015年颁布DMARC协议^[9]为SPF和DKIM协议提供反馈机制。2016年IETF提出基于DNSSEC的DANE与DMARC协议^[10-11]互相补充, 确保安全扩展协议检测结果的正确性及通信双方信道安全。2017年9月公安部、工信部、国家保密局联合下发公信安[2017]2615号文《党

政机关事业单位和国有企业互联网电子邮件系统安全专项整治行动方案》通知, 通知中明确本行动的具体要求: “依托国家电子外网按照统一标准建设安全电子邮件系统, 实现加密互通”。2017年10月16日, 美国国土安全部要求所有运营.gov邮箱域名的联邦机构必须在90天内实施DMARC“监控”政策。2018年6月公安部针对邮件系统安全发布了公信安[2018]1606号文《党政机关事业单位和国有企业互联网电子邮件系统安全等级保护扩展要求(试行)》, 要求电子邮件应满足《信息安全技术网络安全等级保护基本要求》, 多次提到“应使用国产密码技术”来保障邮件传输和存储中安全性。

为了提供电子邮件系统的安全性, 业界提出并标准化了多种安全扩展协议。本文主要探究电子邮件安全扩展协议部署情况, 由于DMARC是基于SPF和DKIM两大主流电子邮件协议发展而来的, 这里主要介绍DMARC协议相关概念。

DMARC协议综合SPF和DKIM协议检测结果提供3方面安全保障: ①检测发件人身份是否真实; ②检测邮件主要内容是否被篡改; ③将SPF/DKIM验证失败的邮件信息发送到管理员邮箱做进一步审查。DMARC记录中包含3个重要参数, p 表示邮件接收者策略, 其中, “none”表示域名所有者要求接收者接收所有的邮件; “quarantine”表示域名所有者希望邮件接收者将DMARC验证失败的邮件标记为可疑的; “reject”表示域名所有者希望邮件接收者拒绝接收DMARC验证失败的邮件, 如表1所示。

表1 DMARC记录重要配置参数及含义
Table 1 The important configuration parameters and meanings of DMARC records

参数	含义
v	DMARC版本号
p	接收邮件策略
ri	报告发送时间间隔
rua	发送综合反馈的邮件地址
ruf	发送消息详细故障信息的邮件地址

DMARC记录相较于SPF和DKIM协议的最大不同在于, 它会将发送域的IP检测结果发送给相应的发件域管理员邮箱, 其中 rua 表示发送

综合反馈的邮件地址, ruf 表示发送消息详细故障信息的邮件地址。

2.2 电子邮件安全扩展协议应用研究

Durumeric 等^[12]通过分析 2015 年 4 月 Gmail 公开的传入电子邮件头部信息, 发现通过 SPF 和 DKIM 协议检测可以验证 94% 的传入电子邮件身份信息。同时指出, 在 Alexa Top1 Million 的邮箱域名中, 只有 47% 的域名部署 SPF 策略, 1% 的域名提供 DMARC 策略; 并且在指定了 SPF 的邮箱域中, 有 29% 的域指定的 IP 地址范围过宽(覆盖了数万个地址)。Hu 等^[13]进一步测量了邮件安全扩展协议部署的有效性, 并指出到 2017 年 45.0% 的域部署了有效的 SPF 协议, 仅有 4.6% 的域部署了有效的 DMARC 协议。同时检测了电子邮件服务提供商如何检测和处理仿冒电子邮件, 结果指出大部分电子邮件服务提供商仍允许仿冒的邮件进入收件箱并且没有任何警告信息。Hu 等^[14]随后召集了 9 个电子邮件服务提供商管理员, 采访他们部署 DMARC 协议的情况, 部署该协议带来的价值以及为什么现在部署 DMARC 协议的域很少。结果指出, DKIM 协议自身技术缺陷、实际部署 DMARC 成本高、风险大, 是很多域未部署 DMARC 协议的主要原因。

然而, 此类研究主要是针对 Alexa Top1 Million 的邮箱域名, 缺乏针对特定企业类型尤其是邮件涉密率较高的政府机构类的电子邮件服务提供商的安全扩展协议采用率的研究。针对电子邮件服务提供商处理安全扩展协议失败的电子邮件的策略的研究, 集中在域仿冒的情况下模拟端到端的网络钓鱼实验, 分析的仿冒策略范围较窄, 缺乏控制单一变量分析不同仿冒策略对邮件服务提供商过滤欺诈邮件效果的影响。

3 电子邮件安全扩展协议测量方法

电子邮件安全扩展协议部署现状分析包括两部分: 测量安全扩展协议部署率和验证部署效果。为了获得部署率, 首先需要获取邮箱域名, 本文首先采用主被动协同的方法采集邮箱域名, 标记邮箱域名属性信息; 然后在获得部署率的基础上, 为了研究电子邮件安全扩展协议部署策略, 本文构建属性图, 关联邮箱域名及属性信息, 提取属性特征, 标记测量结果; 最后, 为了对部

署实际效果进行验证, 本文基于域仿冒特征, 构建欺诈邮件评估模型, 从邮件内容仿冒、域仿冒、信头仿冒 3 个维度对部署效果进行了分析。

3.1 主被动结合的邮箱域名精确标识

邮箱域名标识主要分为两个部分, 分别为主被动协同获取邮箱域名, 以及构建属性图关联邮箱域名精确标识邮箱域名属性信息。属性图 $G=(V, E, \lambda)$ 是有向多重图, 其中, V 表示节点, $E \subseteq (V \times V)$ 是一组边, $\lambda: V \cup E \rightarrow \Sigma$ 表示点与边的一组关系, 如图 1 所示, 具体方法如下。

1) 获取原始数据。通过网络爬虫和被动流量采集获得原始数据 `origin_data`, 使用蓝色长方形标记。网络爬虫的原始数据主要包括网站(企业名称 `company_name`、URL、类别 `company_type` 等, 被动流量采集的原始数据为一个数据包 P_k 。

2) 提取有效数据。企业邮箱通常以网站 `hostname` 后缀或联系人邮箱域名后缀作为邮箱域名, 因此从 `origin_data` 中提取 `hostname`、`connect_email` 等关键字段, 从被动流量数据包 P_k 中提取服务器标识(SNI, `server name indication`)、`from`、`to` 字段。在属性图中使用绿色六边形表示。

3) 标识邮箱域名。解析有效数据中 `hostname` 后缀、`@` 后缀等, 获取可能的邮箱域名集合 D_p , 对集合 D_p 中的元素做 MX 记录验证, 若验证通过, 则将该元素(邮箱域名)加入邮箱域名集合 D_o 中(主动爬虫得到的邮箱域名加入集合 A 中, 被动流量采集得到的邮箱域名加入集合 P 中)。在属性图中使用红色圆形表示邮箱域名, 并连接有效数据与邮箱域名节点生成关系 $\lambda(\text{domain})$ 。

4) 标识邮箱域名属性。邮箱域名属性标识分两步进行, 首先标识主动爬虫得到的邮箱域名集合 A 的属性, 可以直接从 `origin_data` 中获得企业类别 `company_type`、企业名称 `company_name` 等属性, 在属性图中使用紫色四边形表示邮箱域名属性信息, 与邮箱域名节点建立关系 $\lambda(\text{property})$ 。然后标识被动流量采集的邮箱域名集合 P , 由于被动流量中不能直接得到企业类别 `company_type`、企业名称 `company_name` 信息, 通过集合 A 的属性标识分析关系 $\lambda(\text{property})$, 提取不同类别、不同国家域名后缀特征, 通过 ICP 备案验证, 若验证成功, 则在集合 P 中建立 λ 。

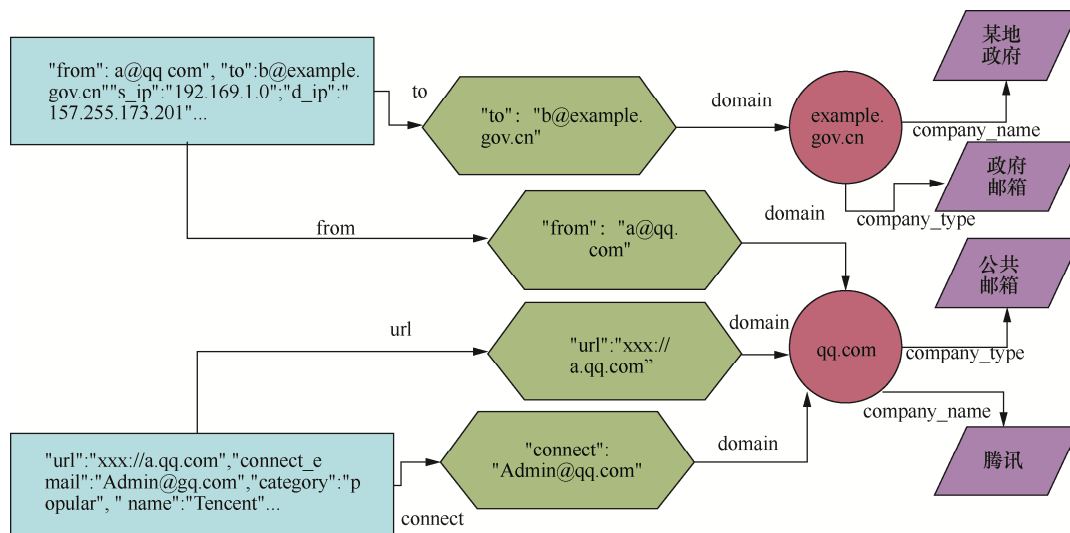


图1 邮箱域名标识属性图(样例1)

Figure 1 Property graph of e-mail domain identification (sample 1)

5) 融合归一。对属性图中邮箱域名节点做相似度计算 $\text{sim}(D_x, D_y)$ ，关联 $\text{sim}(D_x, D_y) \leq K$ 的邮箱域名节点，将属性图中主动获取的邮箱域名集合 A 和被动获取的邮箱域名集合 P 融合为集合 $D = A \cup P = \{x | x \in A \text{ 或 } x \in P\}$ 。

3.2 安全扩展协议部署测量方法

安全扩展协议的测量采用 DNS 解析实现，其测量难点主要由于一些邮件管理员不了解安全扩展协议配置标准，配置时会出现乱码或格式错误，如大小写问题、中英文符号问题等，给识别配置有效性带来干扰。本文提取属性图中邮箱域名

集合 D ，通过查询 DNS 返回结果解析邮箱域名安全扩展协议采用的策略及配置有效性，具体步骤如下。

1) 发送请求。安全扩展协议主要包括 SPF、DKIM、DMARC 这 3 个，因此依次向 DNS 服务器发起 3 个请求 R_{spf} 、 R_{dkim} 、 R_{dmarc} ，查询邮箱域名的 SPF、DKIM 与 DMARC 记录。

2) 标记安全扩展协议查询结果，如图 2 所示。安全扩展协议查询结果标记分 3 方面进行。首先标记 R_{spf} 请求结果， R_{spf} 请求结果包括邮箱域名允许发送的 IP 段 SPF_ip ，协议检测失败推荐

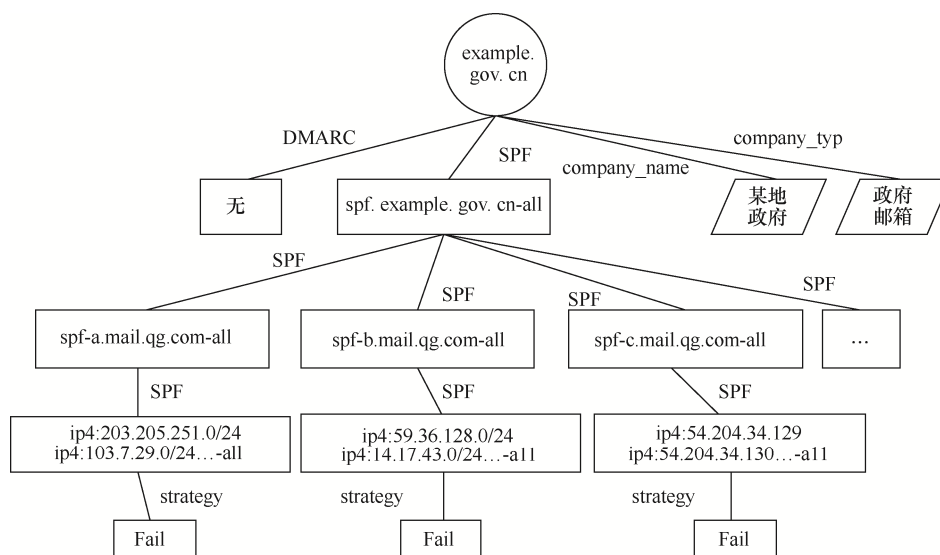


图2 邮箱域名标识属性图(样例2)

Figure 2 Property graph of e-mail domain identification (sample 2)

给接收者使用的策略 SPF_strategy 等, 并与邮箱域名节点 V 建立 $\lambda(\text{SPF})$ 。然后标记 R_{dkim} 请求结果, 请求 R_{dkim} 时要提供邮箱域名的 Selector (选择器) 信息, R_{dkim} 查询结果主要为邮件加密公钥 pk_T , 并与邮箱域名节点 V 建立 $\lambda(\text{DKIM})$ 。最后标记 R_{dmarc} 请求结果, R_{dmarc} 请求结果主要包括协议检测失败推荐给接收者使用的策略 DMARC_strategy , 以及发送报告的管理员邮箱信息 DMARC_report , 并与邮箱域名节点 V 建立 $\lambda(\text{DMARC})$ 。

3) 解析安全扩展协议部署策略。提取 SPF_strategy 和 DMARC_strategy 节点, 解析 SPF 、 DMARC 协议安全扩展协议检测策略 $\text{strategy}(D_i)$ 。

3.3 端到端的电子邮件仿冒策略

为了检测不同安全扩展协议对电子邮件的发送结果的影响, 本文依照不同维度, 设计了三大类电子邮件仿冒策略, 包括邮件内容仿冒、域仿冒、信头仿冒。为了增强实验结果的说服力, 每种策略都会发送若干封正常邮件做对比实验。

(1) 邮件内容仿冒

电子邮件内容仿冒主要包括空白邮件 blank_email 、垃圾邮件内容 spam 、钓鱼 URL 及病毒附件 virus_attach 。由于电子邮件内容会影响垃圾邮件过滤器处理电子邮件的方式。不同的关键字对垃圾邮件过滤器的权重可能不同, 从而影响测量结果, 因此针对邮件内容仿冒的每种形式分别发送 5 封邮件做实验, 保证实验的准确率。

(2) 域仿冒

域仿冒策略是发件人通过远程访问被攻击邮

箱的 MX 记录指向的邮件服务器, 无须账户密码验证, 直接仿冒 mail from 字段的发件人地址, 模仿受信实体, 将仿真邮件发送给接收者邮箱, 进行网络钓鱼, 如图 3 所示。域仿冒策略的发件域选择主要采用以下 7 种形式: ① 没有 SPF 、 DMARC 记录; ② SPF 策略为“soft fail”、无 DMARC 记录; ③ SPF 策略为“soft fail”、 DMARC 记录为“none”; ④ SPF 策略为“soft fail”、 DMARC 记录为“reject”; ⑤ SPF 策略为“fail”、无 DMARC 记录; ⑥ SPF 策略为“fail”、 DMARC 记录为“none”; ⑦ SPF 策略为“fail”、 DMARC 记录为“reject”。

(3) 信头仿冒

信头仿冒策略是发件人通过账号密码登录自己的邮箱, 仿冒邮件头部 from 字段, 将仿真邮件发送给接收者邮箱, 进行网络钓鱼, 如图 4 所示。由于接收人在接收邮件时只能看到信头的发件人, 而一般不会刻意检查邮件头部信息, 也就无法了解真实的发件人, 因而该手段也是黑客进行网络钓鱼的一个策略。在设计欺骗邮件时, 信头仿冒主要采取了两种形式: 同域 (如 $\text{a}@163.\text{com}$ 仿冒为 $\text{b}@163.\text{com}$) 和不同域 (如 $\text{a}@163.\text{com}$ 仿冒为 $\text{b}@qq.\text{com}$)。

4 电子邮件安全扩展协议测量结果

4.1 电子邮件安全扩展协议采用率

在政府机构中电子邮件不仅承载着一种沟通的需要, 更多的时候是作为一种发布告知及传达国家机密消息的工具来使用, 本节针对目前全球政府机构电子邮件安全扩展协议采用率做全面测

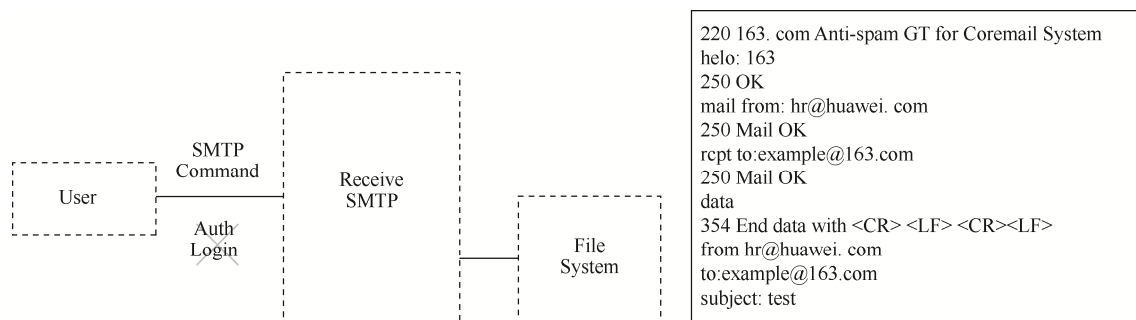


图 3 域仿冒发送电子邮件原理
Figure 3 The principle of sending e-mail by domain phishing

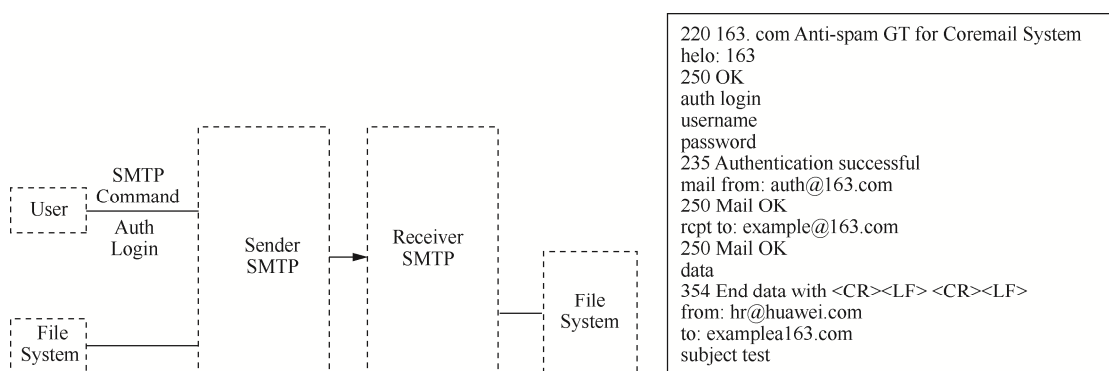


图4 信头仿冒发送电子邮件原理
Figure 4 The principle of sending e-mail by letterhead phishing

量,分析电子邮件安全扩展协议在政府机构中的部署现状,为电子邮件提供商检测欺诈邮件策略提供背景。

4.1.1 数据集情况

基于多通道采集技术、主被动结合,提取属性图中邮箱类别标识为政府的邮箱域名节点。这些节点的数据分别来自:① Alexa 网站,首先按类别获得各国政府网站域名,再从 Alexa 排名前 100 万的域名中,依据域名后缀特征提取各国政府域名(中国 .gov.cn, 美国 .gov、.mil, 日本 .go.jp、.lg.jp, 俄罗斯 .gov.ru、.mid.ru、.mil.ru);② 各国政府官网,经过邮件交换(MX, mail exchanger)记录筛选,统计到的数据情况如表 2 所示。

表2 邮箱域名个数统计
Table 2 E-mail domain number statistics

国家	邮箱域名个数/个
中国	750
美国	2 392
日本	360
俄罗斯	164

4.1.2 测量结果

测量过程中,由于 DKIM 记录需要提供每个域的选择器信息,而选择器信息仅出现在电子邮件头部的 DKIM 签名中,而非公共信息,因此本文不做大规模的 DKIM 测量。测量过程通过属性图标识邮箱域名安全扩展协议,分析各国政府机构是否配置安全扩展协议、采用的策略及配置有

效性。

(1) SPF 协议部署情况(如表 3 所示)

SPF 协议部署率方面,中国政府机构的邮件系统 SPF 协议部署率明显偏低,仅有 40%的域部署了 SPF 记录,其他国家的 SPF 记录配置率达到了 70%以上,日本的政府邮件系统的 SPF 记录配置率达到了 93.33%,配置率较高。

有效性方面,无效的 SPF 记录一方面是管理员配置错误,如将 SPF 记录配置为 DKIM 或 DMARC 记录,或未标明 SPF 记录版本号,或者 SPF 记录配置为乱码;另一方面是因为管理员配置时未采用“严格”策略,使接收者仍能接收到 SPF 检测失败的邮件。实验结果显示,美国政府邮件系统的无效 SPF 记录约 7.68%,部署有效性较弱。

SPF 记录配置策略方面,中国的政府邮件系统都采取拒绝和标记策略,正确配置了 SPF 策略的邮件系统都是严格有效的。美国的政府邮件系统 SPF 记录配置上有 10.17%采用了“?”的形式,这表示如果发送者的邮件服务器 IP 地址未命中 mechanism 中标记的允许发送的 IP 地址范围,也会送达收件人的邮箱中,这同样存在安全隐患。

(2) DMARC 协议部署情况(如表 4 所示)

总体上看,各国的 DMARC 记录配置率都很低,尤其是日本的所有一级行政区政府邮件系统都没配置 DMARC 记录。美国和俄罗斯的 DMARC 协议部署率较其他国家相对高些,然而,部署率也不到 30%,DMARC 协议的部署需要引起各国邮件服务管理员的高度重视。

表 3 各国政府机构 SPF 协议配置策略
Table 3 SPF strategy of government strategies in different countries

国家	有 SPF 记录	SPF 记录无效	SPF 策略为“-”	SPF 策略为“~”	SPF 策略为“?”	SPF 策略为“+”
中国	40.00%	3.23%	55.64%	44.36%	0.00%	0.00%
美国	74.72%	7.68%	44.07%	45.49%	10.17%	0.27%
日本	93.33%	1.19%	58.01%	41.39%	0.60%	0.00%
俄罗斯	72.56%	2.52%	42.27%	54.64%	2.06%	1.03%

表 4 各国政府机构 DMARC 协议配置策略
Table 4 DMARC strategy of government strategies in different countries

国家	有 DMARC 记录	DMARC 记录无效	DMARC 策略为 none	DMARC 策略为 quarantine	DMARC 策略为 reject
中国	3.20%	0.00%	95.83%	0.00%	4.17%
美国	26.66%	9.58%	38.85%	6.83%	54.30%
日本	5.83%	28.57%	75.00%	0.00%	25.00%
俄罗斯	21.34%	20.00%	41.18%	14.71%	44.12%

有效性方面,在配置了 DMARC 记录的邮箱域名中,有 9.28%的域名将 DMARC 记录内容配置成 SPF 记录内容,另有 2.90%格式错误,这些无效的 DMARC 记录占比非常高。

DMARC 记录配置策略方面,目前配置了 DMARC 策略的邮件系统有 40.71%的邮箱域名采用严格有效策略,绝大部分采用的是若 SPF、DKIM 检测失败,仍然将该邮件发到接收者邮箱,这并不代表该配置是不安全的,因为只要配置了 rua 和 ruf 字段,管理员仍能看到这部分 SPF 或 DKIM 检测失败的发送域和 IP。这样配置是因为担心 SPF 或 DKIM 错误检测的正常邮件退回,若反馈到管理员邮箱的某个 IP 地址超过一定次数,并确定该 IP 发送的邮件均为垃圾或钓鱼邮件,再将其策略设置为“reject”。

4.2 端到端的电子邮件仿冒实验

电子邮件是网络钓鱼和社会工程攻击的主要目标。然而,根据 4.1 节电子邮件安全扩展协议采用率的测量结果显示,邮件服务提供商 SPF 和 DMARC 的采用率较低。另外,即使部署了安全扩展协议,也有很多未采用严格有效的策略。对于电子邮件服务提供商来说,可靠地验证所有传入电子邮件仍然是挑战。下文将模拟端到端的电子邮件欺骗实验以检测当电子邮件欺诈检测失败时,电子邮件服务提供商会采取哪些措施应对这些攻击,如何在电子邮件传递和安全性之间进行

权衡。

4.2.1 数据集情况

本文选用常用电子邮箱进行端到端的电子邮件仿冒实验,有 3 个原因:① 据统计,超过 10 亿用户在注册私人电子邮箱时首选公共电子邮件服务提供商,如 Yahoo Mail 和 Gmail,它们的安全策略和设计选择会影响更多的人;② 很多政府机构没有专门的邮件服务管理员搭建管理自己的邮件系统,因而,普遍选择租用公共邮件服务提供商的邮件服务器或直接使用公共邮箱收发邮件;③ 公共电子邮件服务对外开放,使用户可以创建一个账户作为接收者,从接收端搜集数据分析。依据 OnlineEmailVerify 提供的全球 Top100 的常用邮箱列表,成功注册其中的 38 个电子邮箱账户。实验中,将这 38 个邮箱账户作为电子邮件接收端,建立了一个实验服务器,将不同策略仿冒的电子邮件发送到接收者账户,如图 5 所示。该服务器运行 Postfix 邮件服务,使 SMTP 直接与目标邮件服务器进行交互。通过登录这些邮箱账户观察仿冒电子邮件接收结果,推断目标电子邮件服务提供商内部的决策过程。

4.2.2 实验结果

表 5 展示了邮件服务提供商针对欺诈邮件入侵的响应结果。其中○代表正常接收;∞代表该邮件进入垃圾邮箱;△代表标记代发;●代表拒绝接收。

这表明电子邮件服务提供商使用的内容过滤机制为了避免过滤掉正常邮件,涉及的垃圾邮件特征词较少,或最终过滤参数值较低,造成收件箱收到大量垃圾邮件。

(2) 域仿冒

使用域仿冒攻击电子邮件服务系统时,SPF 策略为“fail”,DMARC 策略为“reject”时,邮件服务提供商大多启动响应系统拒绝接收仿冒邮件。然而当域仿冒没有使用严格的安全扩展协议策略时,有 9 个邮箱能正常接收到这些仿冒邮件且没有任何标记。而若发件域未配置任何安全扩展协议时,有 10 个邮箱可以正常接收到这些邮件,还有 6 个邮箱会将其放入垃圾邮箱中。因而攻击者非常喜欢仿冒未配置任何安全扩展协议的邮箱域名作为仿冒的发件域发送欺诈邮件,以跳过接收端邮件服务提供商的欺诈检测机制。实验过程中,还查看了这些邮箱的域仿冒邮件的头部信息,发现大部分有 Authentication-Results 字段,并且对该邮件的检测结果为欺诈检测失败或给予标记,然而在该邮件信头或信体并未发现任何提示信息。

(3) 信头仿冒

使用信头仿冒攻击电子邮件服务系统时,同域、不同域仿冒信头,收件邮箱响应策略基本相同。并且这些仿冒邮件的接收率明显高于使用域仿冒策略的接收率,大部分邮件服务提供商采取将这些仿冒邮件放入垃圾邮箱或标记提醒的策略。

4.2.3 结果分析

本文实验有 4 个关键发现。首先,大多数流行的电子邮件服务提供商具有检测仿冒邮件的功能,但即使接收者执行了所有安全扩展协议检查(SPF、DKIM、DMARC),仍然有允许仿冒的电子邮件进入用户收件箱。其次,通过接收到的电子邮件头部信息发现即使电子邮件服务提供商检测到仿冒邮件,且发件域配置了严格有效的策略,接收域也未必按照发件域给定的策略拒绝接收电子邮件或对域检测失败的电子邮件进行标记。然后,就发送恶意邮件内容而言,大部分邮件服务提供商为了防止过滤掉正常邮件,设置的基于内容的垃圾邮件过滤策略强度很低,恶意邮件可以顺利进入用户收件箱,由用户对恶意邮件进行人

工判别。最后,一旦电子邮件服务提供商收到仿冒的电子邮件,大多数提供商在用户收件箱界面不会向用户发出警告。

5 电子邮件安全扩展协议部署建议及对策

尽管电子邮件承载了一些用户最敏感的信息,但传统电子邮件协议最初并不包括对消息真实性的支持。在过去的 20 年中,邮件系统使用多种安全机制(包括 SPF、DKIM 和 DMARC)对传统电子邮件协议进行了改进。尽管这些身份检测协议可以识别电子邮件发件人真实性及电子邮件内容是否被篡改,但当前配置严格有效的安全扩展协议的邮箱域名较少,不能有效地防护钓鱼邮件攻击。为了强化电子邮件安全扩展协议检测技术的有效性,从而保护用户免遭钓鱼邮件攻击,本节提出以下几个电子邮件服务提供商防护建议。

1) 发布有效的 SPF 协议且 SPF 协议采用严格策略。结合相关 RFC 文件,部署 SPF 协议时建议使用“fail”或“softfail”策略,保证收到仿冒邮件时接收方可以有效过滤掉这些恶意邮件或对该邮件标注代发以警示接收者。

2) 发布有效的 DMARC 协议且标明安全扩展协议检测失败后通知警告的管理员邮箱。将仿冒邮件的发件人及使用的 IP 地址发送到管理员邮箱,以便管理员采取有效措施应对仿冒邮件,将恶意发件人及 IP 地址加入黑名单,也可降低其他邮箱接收恶意邮件的概率。

3) 严格检查接入电子邮件的发件人身份信息。接收域的 SMTP 服务器实时监控和安全扫描技术对接入电子邮件进行安全检测,确定电子邮件身份来源,发现电子邮件异常情况,识别可能发生的攻击行为。严格采用发件域指定的策略过滤并提示电子邮件安全扩展协议检测失败的电子邮件,及时过滤掉仿冒邮件并向发件域管理员邮箱发送安全检测报告,将可疑 IP 列入黑名单。

4) 收件箱界面给予用户安全警告信息。当接收域邮件服务器收到疑似欺诈邮件后,电子邮件服务提供商对该电子邮件应采取显著标记,警告收件账户发件人使用的发件地址与真实发件地址不匹配,安全扩展协议检测失败,根据关键词匹配该邮件为垃圾邮件或钓鱼邮件的概率是多少

等, 这样, 接收者可以及时发现并自我判断该邮件是否为仿冒邮件, 以减少被欺骗的可能。

6 结束语

尽管电子邮件中承载了大量用户敏感信息, 然而, 按照最初的设想, SMTP 协议 (负责在邮件服务器之间中继邮件的协议) 不对发送方进行欺诈检测, 相反, 邮件服务提供商通过一些扩展协议 (SPF、DKIM、DMARC 等) 支持这些功能, 确保邮件传输通信双方的真实性。本文基于主被动协同的方法, 通过构建属性图多表征融合标识邮箱域名, 分析政府机构电子邮件服务提供商安全扩展协议的部署现状及仿冒邮件进入用户邮箱时, 各邮箱给予的安全防护警告策略。实验结果表明, 目前各国政府机构安全扩展协议采用率较低, 尤其是 DMARC 协议, 当仿冒的电子邮件进入用户邮箱时, 大多数电子邮件服务提供商缺少必要的警告机制来通知用户。希望通过阐述安全邮件面临的主要挑战, 邮件服务提供商管理员可以采用有效的邮件服务保护措施, 提升邮件服务系统部署安全性, 以保护用户传输电子邮件的真实性, 免受网络钓鱼攻击。

参考文献:

- [1] POSTEL J B. Request for comments: number 821 simple mail transfer protocol[M]. RFC Editor, 1982.
- [2] FOSTER I D, LARSON J, MASICH M, et al. Security by any other name: on the effectiveness of provider based email security[C]//ACM Conference on Computer and Communications Security (CCS). 2015: 450-464.
- [3] FREED N, BORENSTEIN N. Request for comments: number 2045 multipurpose internet mail extensions (MIME) part one: format of internet message bodies[M]. RFC Editor, 1996.
- [4] ATKINS D, STALLINGS W, ZIMMERMANN P. Request for comments: number 1991PGP message exchange formats [M]. RFC Editor, 1996.
- [5] RAMSDELL B. Request for comments: number 2633 S/MIME version 3 message specification [M]. RFC Editor, 1999.
- [6] HOFFMAN P E. Request for comments: number 3207 SMTP service extension for secure SMTP over transport layer security[M]. RFC Editor, 2002.
- [7] KITTERMAN S. Request for comments: number 7208 sender policy framework (SPF) for authorizing use of domains in email,

version 1[M]. RFC Editor, 2014.

- [8] KUCHERAWY M, CROCKER D, HANSEN T. Request for comments: number 6376 domain keys identified mail (DKIM) signatures[M]. RFC Editor, 2001.
- [9] KUCHERAWY M, ZWICKY E. Request for comments: number 7489 domain-based message authentication, reporting, and conformance (DMARC)[M]. RFC Editor, 2015.
- [10] 柏宗超, 姚健康, 孔宁. 基于 DANE 的电子邮件安全研究[J]. 计算机系统应用, 2018, 27(7): 73-79.
- BAI Z C, YAO J K, KONG N. Research on email security based on DANE[J]. Computer System Application, 2018, 27(7): 73-79.
- [11] HOFFMAN P, SCHLYTER J. Request for comments: number 6698 the DNS-based authentication of named entities (DANE) transport layer security (TLS) protocol: TLSA [M]. RFC Editor, 2012.
- [12] DURUMERIC Z, ADRIAN D, MIRIAN A, et al. Neither snow nor rain nor MITM: an empirical analysis of email delivery security[C]//Internet Measurement Conference (IMC). 2015: 27-39.
- [13] HU H, WANG G. End-to-end measurements of email spoofing attacks[C]//USENIX Security Symposium (USENIX Security). 2018: 1095-1112.
- [14] HU H, PENG P, WANG G. Towards understanding the adoption of anti-spoofing protocols in email systems[C]//IEEE Secure Development Conference (SecDev). 2018: 94-101.

[作者简介]



尚菁菁 (1995—), 女, 北京人, 中国科学院大学硕士生, 主要研究方向为网络空间安全。



朱宇佳 (1984—), 女, 江苏无锡人, 博士, 中国科学院信息工程研究所高级工程师, 主要研究方向为网络空间测绘、安全态势感知技术。



刘庆云 (1980—), 男, 河北人, 博士, 中国科学院信息工程研究所高级工程师, 主要研究方向为网络空间测绘、网络空间安全。