大规模检测和表征横向网络钓鱼

Grant Ho，加州大学伯克利分校和梭子鱼网络； Asaf Cidon、Barracuda Networks 和哥伦比亚大学；梭子鱼网络公司的 Lior Gavish 和 Marco Schweighauser； Vern Paxson，加州大学伯克利分校和 ICSI； Stefan Savage 和 Geoffrey M. Voelker，加州大学圣地亚哥分校；
大卫瓦格纳，加州大学伯克利分校

https://www.usenix.org/conference/usenixsecurity19/presentation/ho

这篇论文包含在第 28 届 USENIX 安全研讨会论文集中。

2019 年 8 月 14-16 日 • 美国加利福尼亚州圣克拉拉

978-1-939133-06-9

开放获取第 28 届 USENIX 安全研讨会论文集
由USENIX赞助。

大规模检测和表征横向网络钓鱼

何炳强[♢] 阿萨夫●西顿[♢ψ] 利奥尔●加维什[♢] 马可●施维豪瑟[♢] 弗恩●帕克森[♢] Stefan Savage Geoffrey M. Voelker 大卫瓦格纳[ψ]

[♢]梭子鱼网络[†] 加州大学伯 UC San Diego[ψ] 哥伦比亚大学国际计算机科学研究所

摘要

我们基于来自 92 个企业组织的 1.13 亿份员工电子邮件的数据集，首次对横向网络钓鱼攻击进行了大规模表征。在横向网络钓鱼攻击中，对手利用受感染的企业帐户向其他用户发送网络钓鱼电子邮件，同时从隐式信任和被劫持用户帐户中的信息中获益。我们开发了一个分类器，可以发现数百封真实世界的横向网络钓鱼电子邮件，同时每 100 万名员工发送的电子邮件中产生的误报不到四次。利用我们检测到的攻击以及用户报告事件的语料库，我们量化了横向网络钓鱼的规模，确定了攻击者遵循的几种主题内容和收件人定位策略，阐明了攻击者表现出的两种复杂行为，并估计这些攻击的成功率。总的来说， 这些结果扩展了我们对“企业攻击者”的心理模型，并阐明了企业网络钓鱼攻击的现状。

1 介绍

十多年来，安全社区探索了无数防御网络钓鱼攻击的方法。然而，尽管工作时间很长，现代攻击者经常并成功地使用网络钓鱼攻击来危害政府系统、政治人物和跨越各个经济部门的公司。这种类型的攻击每年都在日益突出，已经上升到政府关注的水平，FBI 估计 2013 年 10 月至 2018 年 5 月期间报告的 78,617 起事件造成全球 125 亿美元的经济损失 [12]，美国国土安全部部长宣称网络钓鱼是“由最老练的攻击者发起的最具破坏性的攻击”[39].

总的来说，围绕针对谷歌、RSA 和民主党全国委员会等主要实体的针对性鱼叉式网络钓鱼攻击的高调报道已经捕捉并塑造了我们对企业网络钓鱼攻击的心理模型 [35, 43, 46].在这些具有新闻价值的案例中，以及学术文献中讨论的许多有针对性的鱼叉式网络钓鱼事件 [25, 26, 28]，这些攻击来自外部帐户，这些帐户是由民族国家的对手创建的，他们巧妙地制作或欺骗网络钓鱼帐户的名称和电子邮件地址，以模仿已知的合法用户。然而，近年来，这两个行业的工作[7, 24, 36]和学术界[6, 18, 32, 41] 指出了出现和横向网络钓鱼攻击的增长：一种新形式的网络钓鱼它针对各种各样的组织，已经招致了数十亿美元的财务损失 [12].在横向网络钓鱼攻击中，对手使用受损的企业帐户向一组新的收件人发送网络钓鱼电子邮件。这种攻击被证明特别阴险，因为攻击者自动受益于对被劫持帐户的隐含信任：来自人类收件人和传统电子邮件保护系统的信任。

虽然最近的工作 [10, 15, 18, 19, 41] 提出了几种检测横向网络钓鱼的想法，这些先前的方法要么要求组织拥有复杂的网络监控基础设施，要么产生太多误报而无法实际使用。此外，之前的工作还没有在大规模、可概括的范围内描述这种攻击。例如，最全面的相关工作之一使用了来自一个组织的多年数据集，其中仅包含两次横向网络钓鱼攻击 [18].这种情况留下了许多重要的问题没有答案：我们应该如何考虑这类网络钓鱼的规模、复杂性和成功率？攻击者是否遵循主题策略，这些常见行为能否推动新的或改进的防御措施？攻击者如何利用被劫持账户中的信息，他们的行为对企业网络钓鱼攻击的状态和轨迹有何影响？

在学术界和梭子鱼网络之间的这项联合工作中，我们朝着回答这些开放性问题和大规模理解横向网络钓鱼迈出了第一步。本文旨在探索针对这种新兴威胁的实际防御途径，并为这些网络钓鱼攻击的状态开发准确的心理模型。

首先，我们提出了一个新的分类器，用于检测基于 URL 的横向网络钓鱼电子邮件，并在 92 个企业组织的 1.13 亿封电子邮件数据集上评估我们的方法。虽然网络钓鱼电子邮件内容的动态变化和差异性证明具有挑战性，但我们的方法可以检测到我们数据集中 87.3% 的攻击，同时每 1,000,000 封员工发送的电子邮件产生不到 4 次误报。

其次，将我们检测到的攻击与用户报告的横向网络钓鱼攻击的语料库相结合，我们对真实组织中的横向网络钓鱼进行了首次大规模表征。我们的分析表明，这种攻击是强大而广泛的：数十家组织，从少于 100 名员工的组织到超过 1,000 名员工的组织，都在这一范围内遭受了横向网络钓鱼攻击

几个月;在一组随机抽样的组织中，总共有 14% 在七个月的时间跨度内经历了至少一次横向网络钓鱼事件。此外，我们估计超过 11% 的攻击者成功危害了至少一名额外的员工。尽管我们的真实来源和检测器面临限制其发现隐蔽或针对性攻击的能力的限制，但我们的结果仍然阐明了目前影响许多现实世界组织的突出威胁。

通过检查横向网络钓鱼者的行为，我们探索并量化了四种接收者（受害者）选择策略的流行程度。尽管我们数据集的攻击者以数十到数百个收件人为目标，但这些收件人通常包括与被劫持帐户有某种关系的用户子集（例如，同事或最近的联系人）。此外，我们为数据集的网络钓鱼消息显示的不同级别的内容定制开发了一个分类。我们的分类表明，虽然 7% 的攻击部署了有针对性的消息，但大多数攻击选择了网络钓鱼者可以轻松在多个组织中重复使用的通用内容。特别是，我们观察到横向网络钓鱼者主要依赖两种常见的诱饵：共享文档的借口和关于收件人帐户问题的虚假警告消息。尽管非针对性内容很受欢迎， 我们数据集中近三分之一的攻击者投入了额外的时间和精力来使他们的攻击更具说服力和/或逃避检测；而且，超过 80% 的攻击发生在被劫持账户的正常工作时间。

最终，这项工作产生了两个贡献，扩展了我们对企业网络钓鱼和潜在防御措施的理解。首先，我们提出了一种新颖的检测器，它比之前的工作实现了一个数量级的性能提升，同时以最少的数据要求（仅利用历史电子邮件）运行。其次，通过首次对横向网络钓鱼进行大规模表征，我们揭示了这种新兴攻击的规模和成功率，并阐明了横向网络钓鱼者采用的常见策略。我们的分析阐明了一类普遍存在的企业攻击者，他们的行为并不完全符合有针对性的民族国家攻击或工业间谍活动的策略。尽管如此，这些横向网络钓鱼者仍然在没有新防御措施的情况下取得成功，而且我们数据集的许多攻击者确实表现出一些老练和专注的迹象。

2 背景

在横向网络钓鱼攻击中，攻击者使用受感染但合法的电子邮件帐户向受害者发送网络钓鱼电子邮件。攻击者的目标和恶意负载的选择可以采用多种不同的形式，从感染恶意软件的附件到网络钓鱼 URL，再到虚假的支付请求。我们的工作重点是使用嵌入电子邮件中的恶意 URL 的横向网络钓鱼攻击，这是我们数据集中确定的最常见的利用方法。

清单 1：使用虚假合同文件作为引诱的横向网络钓鱼消息的匿名示例。

```
来自：《爱丽丝》
<alice@company.com>致：“鲍
勃”<bob@company.com> 主题：X公
司（新合约）

新合约

查看文档［此文本链接到网络钓鱼网站］问候，
爱丽丝[签名]
```

清单1 显示了我们研究中横向网络钓鱼攻击的匿名示例。在这次攻击中，网络钓鱼者试图以新合同为幌子诱使收件人点击链接。此外，攻击者还试图通过回复询问电子邮件真实性的收件人来提高欺骗的可信度；他们还通过删除所有网络钓鱼电子邮件的痕迹，主动隐藏在受感染用户的邮箱中。

横向网络钓鱼是一种危险但未被充分研究的攻击，处于网络钓鱼和帐户劫持的交叉点。从广义上讲，网络钓鱼攻击涉及攻击者从任何帐户（被盗用或欺骗）制作欺骗性电子邮件，以诱使受害者执行某些操作。帐户劫持，在行业术语中也称为帐户接管（ATO），涉及使用受感染的帐户进行任何类型的恶意手段（例如，包括垃圾邮件）。虽然之前的工作主要是在较小规模和针对个人账户的情况下检查这些攻击中的每一个，但我们的工作是从企业组织的角度大规模研究这两种攻击的交集。通过这样做，我们扩展了对重要企业威胁、防御这些威胁的途径以及实施这些威胁的攻击者所使用的策略的理解。

2.1 相关工作

检测：大量现有文献提出了多种检测传统网络钓鱼攻击的技术［1,3,13,14,44］，以及更复杂的鱼叉式网络钓鱼攻击［8, 10, 23, 41, 47］.胡等。研究了如何使用社交图指标来检测从受感染帐户发送的恶意电子邮件［19］.他们的方法检测到的被劫持账户的误报率在 20-40% 之间。不幸的是，在实践中，许多组织每天处理数以万计的员工发送的电子邮件，因此 20% 的误报率将导致每天数千次误报。 IdentityMailer，由 Stringhini 等人提出。［41］，通过基于每个用户的时间模式、元数据和文体测量法训练行为模型来检测横向网络钓鱼攻击。如果一封新邮件偏离了员工的行为模式，

他们的系统将其标记为攻击。虽然很有前途，但他们的方法会产生 1-10% 的误报率，鉴于大量良性电子邮件和较低的网络钓鱼基本率，这在实践中是站不住脚的。此外，他们的系统需要为每个员工培训一个行为模型，从而产生昂贵的技术债务以进行大规模操作。

何等。通过对从历史用户登录数据和企业网络流量日志派生的一组特征应用一种新颖的异常检测算法，开发了检测横向鱼叉式网络钓鱼的方法 [18]. 他们的方法检测已知和新发现的攻击，误报率为 0.004%。然而，技术专长较少的组织通常缺乏全面捕获企业网络流量的基础设施，而这是现有方法所需要的。这个技术先决条件引出了一个问题，我们是否可以使用更简约的数据集实际检测横向网络钓鱼攻击：仅企业的历史电子邮件？此外，他们的数据集反映了一家企业在 3.5 年的时间跨度内只经历了两次横向网络钓鱼攻击，这使他们无法在一般范围内描述横向网络钓鱼的性质。

特征：虽然之前的工作表明攻击者经常使用网络钓鱼来破坏帐户，并且攻击者偶尔会从这些被劫持的帐户进行（横向）网络钓鱼，但很少有人深入和大规模地研究横向网络钓鱼的性质。检查来自谷歌数据源的网络钓鱼电子邮件、网页和受感染帐户的样本，一项先前的帐户劫持研究发现，攻击者经常使用这些帐户向帐户的联系人发送网络钓鱼电子邮件 [6]. 然而，他们得出的结论是，自动检测此类攻击具有挑战性。 Onaolapo 等人。研究了攻击者如何处理被劫持的账户 [32]，但他们没有检查横向网络钓鱼。与电子邮件帐户不同，一项针对受感染的 Twitter 帐户的研究发现，感染似乎通过社交网络横向传播。然而，他们的数据集不允许直接观察横向攻击向量本身 [42]，它也没有提供对受损企业帐户领域的见解（鉴于社交媒体的性质）。

未解决的问题和挑战：之前的工作清楚地表明，帐户泄露是一个重大而普遍的问题。该文献还为已实施复杂监控的企业提供了有前途的防御措施。然而，尽管取得了这些进展，但仍有几个关键问题没有得到解决。没有全面监控和技术专长的组织是否有切实可行的方法来防御横向网络钓鱼攻击？横向网络钓鱼者采用哪些常见策略和交易技巧？横向网络钓鱼者如何利用他们对合法帐户的控制，他们的战术复杂性对企业网络钓鱼的状态有何影响？本文通过提出一种新的检测策略和对横向网络钓鱼攻击的大规模表征，朝着回答这些悬而未决的问题迈出了一步。

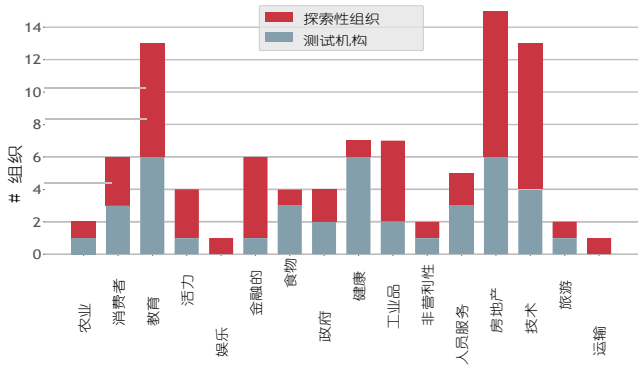

图 1：我们数据集的 52 个探索组织与 40 个测试组织的经济部门细分。

2.2 伦理

在这项工作中，我们的团队由来自学术界和一家大型安全公司的研究人员组成，他们使用历史电子邮件数据集开发了检测技术，并报告了 92 个组织的事件，这些组织是 Barracuda Networks 的活跃客户。这些组织授予 Barracuda 访问其 Office 365 员工邮箱的权限，以研究和开发针对横向网络钓鱼的防御措施。根据 Barracuda 的政策，所有提取的电子邮件都以加密方式存储，客户可以随时选择撤销对其数据的访问权限。

由于数据的敏感性，只有经过授权的梭子鱼员工才能访问这些数据（在标准、严格的访问控制政策下）。没有与任何非梭子鱼员工共享个人身份信息或敏感数据。我们的项目还获得了 Barracuda 的合法批准，梭子鱼获得了客户的许可，可以对数据进行分析和操作。

一旦梭子鱼将一组横向网络钓鱼检测器部署到生产环境中，任何检测到的攻击都会实时报告给客户，以防止经济损失和伤害。

3 数据

我们的数据集包含来自 92 个英语组织的员工发送的电子邮件； 23 个组织来自随机抽样的有横向网络钓鱼报告的企业，从所有组织中随机抽取 69 个。在这些企业中，25 个组织拥有 100 个或更少的用户帐户，34 个组织拥有 101-1000 个帐户，33 个组织拥有超过 1000 个帐户。房地产、科技和教育构成了我们数据集中最常见的三个行业，分别有 15 家、13 家和 13 家企业；人物1 和2 显示经济部门的分布和我们数据集组织的规模，按探索性组织与测试组织细分（§3.2).

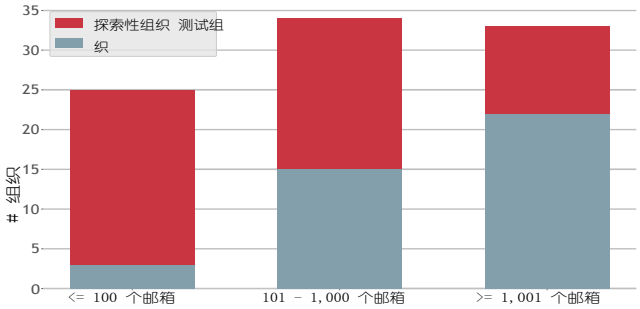

图 2：我们数据集的 52 个探索性组织与 40 个测试组织的组织规模细分。

3.1 图式

我们数据集中的组织使用 Office 365 作为他们的电子邮件提供商。在较高级别，每个电子邮件对象包含：一个唯一的 Office 365 标识符；电子邮件的元数据（SMTP 标头信息），它描述了诸如电子邮件的发送时间戳、收件人、声称的发件人和主题等属性；和电子邮件的正文，即完整 HTML 格式的电子邮件内容。Office 365 的文档描述了每个电子邮件对象的完整架构 [29]. 此外，对于每个组织，我们都有一组经过验证的域：组织已声明其拥有的域。

3.2 数据集大小

我们的数据集包含 113,083,695 封独特的员工发送的电子邮件。为了确保我们的检测技术得到推广（第5.1），我们将数据分为 2018 年 4 月至 6 月期间来自 52 个“探索性组织”的电子邮件训练数据集，以及来自 92 个组织的涵盖 2018 年 7 月至 10 月的测试数据集。我们的测试数据集包含来自 52 个探索性组织的电子邮件（但来自比我们的训练数据集更晚、不相交的时间段），以及来自另外一组保留的 40 个“测试组织”的数据。我们通过在分析任何数据之前执行的随机样本选择了 40 个测试组织。我们的训练数据集有 25,670,264 封电子邮件，我们的测试数据集有 87,413,431 封电子邮件。如图所示，这两组组织涵盖了不同的行业和规模1 和2. 探索组织共有 89,267 个发送或接收电子邮件的用户邮箱，测试组织有 138,752 个邮箱（基于 2018 年 10 月的数据）。[1]

3.3 地面实况

我们的一组横向网络钓鱼电子邮件来自两个来源：(1) 组织的安全管理员向梭子鱼报告的攻击电子邮件，以及用户向其组织或直接向梭子鱼报告的攻击，以及 (2) 电子邮件

[1]由于使用邮件列表和别名，邮箱数量是员工数量的上限。

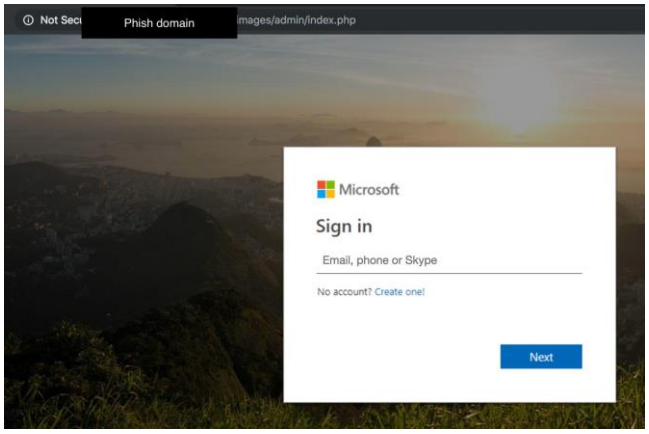

图 3：横向网络钓鱼电子邮件中的网络钓鱼 URL 导致的网页的匿名屏幕截图。

由我们的检测器标记（§4），我们在包含之前对其进行了人工审查和标记。

在高层次上，为了手动将电子邮件标记为网络钓鱼，我们检查了它的邮件内容、Office 365 元数据和 Internet 邮件标头 [33] 以确定电子邮件是否包含网络钓鱼内容，以及电子邮件是否来自受感染的帐户（相对于外部帐户，我们不将其视为横向网络钓鱼）。例如，如果 Office 365 元数据显示电子邮件的副本位于员工的“已发送邮件”文件夹中，或者如果其标头显示电子邮件通过了相应的 SPF 或 DKIM [9] 检查，然后我们认为该电子邮件是横向网络钓鱼。附录 §A.1 详细描述了我们的标签程序。

此外，对于这些横向网络钓鱼电子邮件中的一小部分 URL，梭子鱼的员工在包含 VM 的浏览器中访问了网络钓鱼 URL，以更好地了解攻击的最终目标。为了尽量减少潜在的危害和副作用，这些员工只访问了不包含唯一标识符（即 URL 路径中没有随机字符串或用户/组织信息）的网络钓鱼 URL。为了处理驻留在 URL 缩短域上的任何网络钓鱼 URL，我们使用了 Barracuda 的 URL 扩展 API 之一，其生产服务已经应用于电子邮件 URL，并且仅访问扩展为非副作用 URL 的可疑网络钓鱼链接。我们探索的大多数网络钓鱼 URL 都指向一个 SafeBrowsing 插页式网页，这可能反映了我们对历史电子邮件的使用，而不是用户同时遇到的情况。然而，最近的恶意 URL 不断导致凭证钓鱼网站设计得看起来像合法的 Office 365 登录页面（我们研究的组织使用的电子邮件服务提供商）；数字3 显示了一个网络钓鱼网站的匿名示例。我们的数据集总共包含 1,902 封横向网络钓鱼电子邮件（主题、发件人和发送时间各不相同），由来自 33 个组织的 154 个被劫持的员工帐户发送。其中 1,694 封电子邮件是由用户报告的，其余的仅由我们的检测器发现（§4）；我们的检测器还发现了很多

用户报告的攻击以及（§5）.在用户报告的攻击中，40 封电子邮件（来自 12 个被劫持的帐户）包含虚假电汇或恶意附件，而其余 1,862 封电子邮件使用了恶意 URL。

鉴于这种攻击媒介的普遍存在，我们将检测策略集中在基于 URL 的网络钓鱼上。这种关注意味着我们的分析和检测技术无法反映横向网络钓鱼攻击的全部空间。尽管存在这一限制，我们的数据集的攻击跨越了数十个组织，使我们能够研究一类普遍存在的企业网络钓鱼，这种网络钓鱼本身就构成了重要威胁。

4 检测横向网络钓鱼

采用 Ho 等人定义的横向攻击者威胁模型。［18］，我们专注于由受感染的员工帐户发送的网络钓鱼电子邮件，其中攻击嵌入了恶意 URL 作为漏洞利用（例如，将用户引导至网络钓鱼网页）。我们探索了三种检测横向网络钓鱼攻击的策略，但最终发现其中一种策略几乎检测了所有三种方法识别的所有攻击。在高层次上，这两种不太有效的策略通过查找包含 (1) 罕见 URL 和 (2) 其文本似乎可能用于网络钓鱼的消息（例如，与已知网络钓鱼攻击相似的文本）的电子邮件来检测攻击.因为我们的主要检测策略只检测到其他策略发现的攻击中的两种，同时发现的攻击次数是其他策略的十倍以上，所以我们将对两种不太成功的方法的讨论推迟到我们的扩展技术报告［17］;下面，我们着重详细探讨更有效的策略。在我们的评估中，我们将替代方法发现的另外两种攻击作为我们检测器的漏报。

概述：我们在我们的训练数据集中（2018 年 4 月至 6 月）检查了用户报告的横向网络钓鱼事件，以确定我们可以在我们的检测器中利用的广泛主题和行为。通过发送这些攻击的被劫持帐户（ATO）对这组攻击进行分组，我们发现其中 95% 的 ATO 向 25 个或更多不同的收件人发送了网络钓鱼电子邮件。[2] 这种普遍的行为，以及受诱饵漏洞利用检测框架启发的其他功能想法［18］，为我们的检测策略提供了基础。在本节的其余部分，我们将描述我们的检测器使用的功能、这些功能背后的直觉，以及我们的检测器用于对电子邮件进行分类的机器学习过程。

我们的技术既没有提供一种包罗万象的方法来发现每一次攻击，也没有保证对试图逃避检测的有动机的对手的鲁棒性。但是，我们在部分中显示5 我们的方法在数十个真实世界的组织中发现了数百封横向网络钓鱼电子邮件，同时产生了少量的误报。

[2]为了评估我们方法的普遍性，我们的评估使用了一个保留的数据集，来自较晚的时间范围和新的组织（§5).

特征：我们的检测器提取三组特征。第一组包含两个针对我们之前观察到的流行行为的特征：联系许多收件人。给定一封电子邮件，我们首先从电子邮件的“收件人”、“抄送”和“密件抄送”标头中提取唯一收件人的数量。此外，我们计算此电子邮件的收件人集与上个月任何员工发送的电子邮件中最接近的历史收件人集的 Jaccard 相似度。我们将后一种（相似性）特征称为电子邮件的收件人可能性得分。

接下来的两组特征借鉴了 Ho 等人提出的诱饵利用网络钓鱼框架。［18].该框架假定网络钓鱼电子邮件包含两个必要的组件：“诱饵”，它说服受害者相信网络钓鱼电子邮件并执行某些操作；和“利用”：受害者应该执行的恶意行为。他们的工作发现，使用同时针对这两个组件的功能可以显着提高检测器的性能。

为了确定一封新电子邮件是否包含潜在的网络钓鱼诱饵，我们的检测器根据电子邮件的文本提取了一个单一的、轻量级的布尔特征。具体来说，梭子鱼为我们提供了一组大约 150 个经常出现在网络钓鱼攻击中的关键字和短语。他们通过从数百个真实世界的网络钓鱼电子邮件（包括外部网络钓鱼和横向网络钓鱼）中提取链接文本并选择在这些攻击中出现最频繁的（规范化）文本来开发这组“网络钓鱼”关键字。从主题上讲，这些可疑的关键字传达了一种号召性用语，诱使收件人单击链接。对于我们的“诱饵”功能，我们提取了一个布尔值，该值指示电子邮件是否包含这些钓鱼关键字中的任何一个。

最后，我们通过提取两个捕获电子邮件是否可能包含漏洞利用的特征来完成检测器的特征集。由于我们的工作重点是基于 URL 的攻击，这组特征反映了电子邮件是否包含潜在危险的 URL。

首先，对于每封电子邮件，我们提取一个全局 URL 信誉特征，该特征量化电子邮件包含的最罕见的 URL。给定一封电子邮件，我们从电子邮件正文中提取所有 URL，并忽略属于以下两类的 URL：我们排除其域在组织的已验证域列表中列出的所有 URL（§3.1），并且我们还排除了所有显示的超链接文本与超链接的基础目标 URL 完全匹配的 URL。例如，在列表中1在 的攻击中，钓鱼超链接显示的文本是“单击此处”，与超链接的目标（钓鱼站点）不匹配，因此我们的程序将保留此 URL。相比之下，Alice 的签名来自 Listing1 可能包含指向她的个人网站的链接，例如 www.alice.com；我们的程序会忽略这个 URL，因为显示的文本www.alice.com 匹配超链接的目的地。

后一种过滤标准假设网络钓鱼 URL 会尝试混淆自身，并且不会直接向用户显示真正的底层目标。在这些过滤步骤之后，我们通过

将每个剩余的 URL 映射到其注册域，然后在 Cisco Umbrella 前 100 万个站点上查找每个域的排名 [20];[3] 对于任何未列出的域，我们为其分配默认排名 1000 万。我们区别对待两种特殊情况。对于较短域上的 URL，我们的检测器会尝试递归地将短链接解析到其最终目的地。如果此解析成功，我们将使用最终 URL 域的全球排名；否则，我们将 URL 视为来自未排名的域（1000 万）。对于内容托管网站（例如 Google Drive 或 Sharepoint）上的 URL，如果不获取内容并对其进行分析（该操作有几个实际障碍），我们没有好的方法来确定其可疑性。因此，我们将内容托管网站上的所有 URL 视为驻留在未排名的域中。在对每个 URL 的域进行排名后，我们将电子邮件的全球 URL 信誉功能设置为其 URL 中最差（最高）的域排名。直觉上，我们预计网络钓鱼者很少会在热门网站上托管网络钓鱼页面，因此较高的全球 URL 信誉表示更可疑的电子邮件。 原则上，有动机的对手可以规避此功能；例如，如果对手可以破坏组织的一个已验证域，他们可以从这个被破坏的站点托管他们的网络钓鱼 URL 并避免准确排名。但是，我们在用户报告的横向网络钓鱼集中没有发现此类实例。此外，由于本文的目标是开始探索实用的检测技术，并开发大量横向网络钓鱼事件供我们分析，因此此功能足以满足我们的需求。

除了这个全球声誉指标外，我们还提取了一个本地指标，该指标描述了 URL 相对于组织员工通常发送的 URL 域的稀有性。给定电子邮件中嵌入的一组 URL，我们将每个 URL 映射到其完全限定的域名（FQDN），并计算从上个月起至少有一封员工发送的电子邮件包含 FQDN 上的 URL 的天数。然后我们在所有电子邮件的 URL 中取最小值；我们将这个最小值称为本地 URL 信誉特征。直觉上，可疑的 URL 将具有低全球声誉和低本地声誉。然而，我们的评估（§ 5.2）发现此本地 URL 信誉功能增加的价值很小：本地 URL 信誉值较低的 URL 几乎总是具有较低的全局 URL 信誉值，反之亦然。

分类：为了将电子邮件标记为网络钓鱼，我们训练了一个随机森林分类器 [45] 具有上述特点。为了训练我们的分类器，我们在训练数据集中获取所有用户报告的横向网络钓鱼电子邮件，并将它们与一组可能是良性的电子邮件结合起来。我们通过随机抽取训练窗口中未被报告为网络钓鱼的电子邮件的子集来生成这组“良性”电子邮件；我们为每个人抽取 200 封这样的良性电子邮件攻击电子邮件以形成我们用于训练的良性电子邮件集。按照标准的机器学习实践，我们选择了分类器的超参数和使用此训练数据的交叉验证的精确下采样率（200:1）。附录 A.2 更详细地描述了我们的培训程序。一旦我们有了一个训练有素的分类器，给定一封新电子邮件，我们的检测器就会提取它的特征，将特征输入这个分类器，并输出分类器的决定。

[3] 我们使用 2018 年 3 月上旬获取的列表进行特征提取，但实际上，可以使用不断更新的列表。

5 评估

在本节中，我们评估我们的横向网络钓鱼检测器。我们首先描述了我们的测试方法，然后展示了检测器对来自 90 多个组织的数百万封电子邮件的执行情况。总的来说，我们的检测器检测率高，误报率低，检测到许多新的攻击。

5.1 方法

建立通用性：如前面部分所述3.2，我们将数据集分成两个不相交的部分：一个训练数据集，由 2018 年 4 月至 6 月期间来自 52 个探索性组织的电子邮件组成，另一个是 2018 年 7 月至 10 月期间来自 92 个企业的测试数据集；在 §5.2，我们表明，如果我们的测试数据集仅包含来自 40 个隐瞒测试组织的电子邮件，我们的检测器性能保持不变。鉴于这两个数据集，我们首先训练我们的分类器并通过对我们的训练数据集进行交叉验证来调整其超参数（附录A.2）.接下来，为了计算我们的评估结果，我们在保留的测试数据集的每个月运行我们的检测器。为了模拟生产中的分类器，我们遵循标准的机器学习实践，并使用连续学习程序每月更新我们的检测器 [38].即，在每个月底，我们将之前所有月份的用户报告和检测器发现的网络钓鱼电子邮件汇总到一组新的网络钓鱼“训练”数据中；并且，我们将我们的原始随机抽样良性电子邮件集与我们的检测器从所有前几个月的误报汇总起来，形成一个新的良性“训练”数据集。然后，我们在这个聚合的训练数据集上训练了一个新模型，并使用这个更新后的模型对下个月的数据进行分类。然而，为了确保我们从训练数据集中获得的任何调整或知识不会偏向或过度拟合我们的分类器，我们在对测试数据集进行评估期间没有改变任何模型的超参数或特征。

我们的评估在训练和测试数据集之间进行时间分割，以及将随机保留的组织的新数据引入测试数据集，遵循推荐这种方法而不是随机交叉验证评估的最佳实践 [2, 31, 34].完全随机的评估（例如，交叉验证）存在对未来数据进行训练和对过去进行测试的风险，这可能会导致我们高估检测器的有效性。相比之下，

公制	训练 2018 年 4 月 – 6 月	测试 2018 年 7 月至 10 月
组织	52探索性	52探索性 + 40 测试
检测到已知攻击	34	47
检测到新的攻击	28	49
未命中 (FN)	8	14
检测率	88.6%	87.3%
电子邮件总数	25,670,264	87,413,431
误报 (FP)	136	316
误报率	0.00053%	0.00036%
精确	31.3%	23.3%

表 1：我们检测器的评估结果。 “已知检测”显示了我们的检测器识别的事件数量，并且还由组织的员工报告。 “检测到的新攻击”显示了我们的检测器发现但没有人报告的事件数量。 “未命中攻击 (FN)”显示了所有由用户报告或由我们的任何检测策略发现的事件，但我们的检测器将其标记为良性（假阴性）。在我们的检测器遗漏的 22 起事件中，有 12 起是基于附件的攻击，这是一种我们的检测器明确不针对的威胁模型，但为了完整性，我们将其包含在我们的 FN 和检测率结果中。

我们的方法使用“未来”时间段的新数据评估我们的检测器，并引入了 40 个新组织，我们的检测器在训练期间都没有发现这些组织；这也反映了探测器在实践中的运作方式。

警报指标（事件）：我们有多种选择来对检测器的警报生成过程进行建模（即，我们如何计算不同的攻击）。例如，我们可以根据检测器正确标记的独特电子邮件数量来评估检测器的性能。或者，我们可以根据检测器标记为已泄露的不同员工帐户的数量来衡量检测器的性能（对每个帐户生成一个警报并抑制其余警报的检测器建模）。最终，我们选择了一个实践中常用的概念，即事件，它对应于一个唯一的（主题，发件人电子邮件地址）对。在这个粒度上，我们的检测器的警报生成模型为每个唯一的（主题，发件人）对生成一个警报。该指标避免了有偏见的评估数字，这些数字过分强调在一次攻击期间生成许多相同电子邮件的妥协事件。例如，如果有两个事件，其中一个事件分别向一个收件人生成一百封电子邮件， 另一个生成一封电子邮件给 100 个收件人，如果我们计算电子邮件级别的攻击，检测器在一百封电子邮件事件中的性能将主导结果。

总的来说，我们的训练数据集包含 40 个来自用户报告的真实来源的横向网络钓鱼事件，我们的测试数据集包含 61 个用户报告的事件。我们的检测器发现了另外 77 起未报告的事件（表第 2 行1).

5.2 检测结果

桌子1 总结了我们检测器的性能指标。我们使用术语检测率来指代我们的检测器发现的横向网络钓鱼事件，除以我们数据集中所有已知的攻击事件（即，任何用户报告的事件和我们尝试过的任何检测技术发现的任何事件）。为了完整起见，我们将 12 个基于附件的事件包括在我们的漏报率和检测率计算中，我们的检测器显然错过了这些事件，因为我们设计它是为了捕获基于 URL 的横向网络钓鱼。此外，我们还包括作为假阴性的 2 次训练事件，这些事件是我们不太成功的检测器识别出来的 [17]；这两种替代策略在测试数据集中没有发现任何新的攻击。因此，检测率反映了一种尽力而为的评估，它可能高估了我们检测器的真实阳性率，因为我们有一个不完美的基本事实，无法解释用户未报告的狭隘目标攻击。精度等于我们的检测器生成的攻击警报（事件）的百分比除以我们的检测器生成的警报总数（攻击加上误报）。

培训和调整：在培训数据集上，我们的检测器正确识别了 70 起横向网络钓鱼事件中的 62 起 (88.6%)，同时产生了总共 62 起误报（在 2570 万封员工发送的电子邮件中）。

我们的 PySpark 随机森林分类器公开了每个特征相对重要性的内置估计 [40]，其中每个特征的得分在 0.0–1.0 之间，所有得分的总和为 1.0。基于这些特征权重，我们的模型最重视全局 URL 信誉特征，赋予它 0.42 的权重，以及电子邮件的“收件人数量”特征 (0.34)。相比之下，我们的模型基本上忽略了我们的本地 URL 信誉，为其分配了 0.01 的分数，这可能是因为大多数全球稀有域往往在本地也是稀有的。在其余特征中，收件人似然特征的权重为 0.17，“phishy”关键字特征的权重为 0.06。

测试数据集：我们的检测器在我们的地面实况数据集中的 110 起测试事件 (87.3%) 中正确识别了 96 起横向网络钓鱼事件。此外，我们的检测器还发现了 49 起事件，根据我们的基本事实，这些事件并未被用户报告为网络钓鱼。就其成本而言，我们的检测器在整个测试数据集中产生了 312 个误报（误报率低于 0.00035%，假设未被我们的基本事实识别为攻击的电子邮件是良性的）。在我们的测试数据集中，92 个组织中的 82 个组织在整个四个月的窗口中累积了 10 个或更少的误报，其中 44 个组织在此时间跨度内遇到了零误报。相比之下，在所有四个月中，只有三个组织的误报总数超过 40 次（分别遇到 44 次、66 次和 83 次误报）。 我们的检测器如果只对我们 40 家隐瞒测试机构的数据进行评估，也能达到类似的结果，检测率为 91.0%，精度为 23.1%，误报率为 0.00038%。

偏见和规避：我们的评估数字基于我们拥有的最佳基本事实：所有用户报告的组合

横向网络钓鱼事件（包括我们的威胁模型之外的一些攻击），以及我们尝试过的任何检测技术发现的所有事件（包括与我们的检测器策略正交的两种方法）。这个基本事实存在对联系许多潜在受害者的网络钓鱼电子邮件的偏见，以及用户更容易识别的攻击。此外，由于我们的检测器专注于基于 URL 的漏洞利用，我们的攻击数据集可能低估了非基于 URL 的网络钓鱼攻击的普遍性，这些攻击仅来自我们数据集中用户报告的实例。因此，我们的工作并没有捕捉到横向网络钓鱼攻击的全部空间，例如攻击者针对范围狭窄、选择的一组具有隐秘欺骗性内容的受害者的攻击。相反，鉴于我们的检测器识别了许多已知和未报告的攻击，同时每月仅产生一些误报，我们为未来工作可以扩展的实际检测提供了一个起点。 此外，即使我们的检测器没有捕捉到所有可能的攻击，但我们数据集中的攻击跨越数十个不同的组织，跨越数月的时间框架，这一事实使我们能够阐明许多企业目前面临的一类未被充分研究的攻击。

除了获得更全面的基本事实外，还需要做更多的工作来探索针对潜在逃避攻击的防御措施。攻击者可能会尝试通过针对我们利用的不同特征（例如他们所针对的收件人的组成或数量）来逃避我们的检测器。针对这些逃避攻击中的许多，未来的工作可以利用其他功能和数据，例如用户在电子邮件帐户中采取的操作（例如，侦察操作，例如异常搜索，表明攻击者挖掘目标收件人的帐户以进行攻击）或来自用户帐户登录的信息（例如，Ho 等人提出的检测器使用帐户的登录 IP 地址 [18] 来检测横向网络钓鱼）。同时，未来的工作应该研究攻击者进行哪些逃避攻击在经济上仍然可行。例如，攻击者可以选择只针对少数用户以逃避我们的检测器；但即使这种逃避成功了，欺骗接收者的转换率也可能非常低，以至于攻击最终无法损害经济上可行的受害者数量。事实上，正如我们在下一节中探讨的那样（§6），在我们的数据集中捕获的攻击者已经参与了一系列不同的行为，包括一些复杂的手动操作形式，以提高攻击的成功率。

6 表征横向网络钓鱼

在本节中，我们使用整个数据集（训练和测试）中的所有已知攻击对真实世界的横向网络钓鱼进行分析。在七个月的时间跨度内，共有 33 个组织经历了横向网络钓鱼攻击，其中大多数受感染的组织都经历了多起事件。通过检查攻击的主题消息内容和收件人定位策略，我们

规模与成功	
# 不同的钓鱼邮件	1,902
# 事件	180
# ATO	154
# 个组织发生了 1 次以上事件	33
# 网络钓鱼收件人	101,276
成功的 ATO 百分比	11%
# 妥协的员工收据（平均）	542

表 2：我们数据集中横向网络钓鱼攻击的规模和成功的总结（§6.1）.

分析表明，我们数据集中的大多数横向网络钓鱼者不会主动挖掘被劫持帐户的电子邮件来制作个性化的鱼叉式网络钓鱼攻击。相反，这些攻击者以投机取巧的方式运作，并依赖于常见的网络钓鱼内容。这一发现表明，企业网络钓鱼的空间已经超出了其与复杂的 APT 和民族国家对手的历史联系。

与此同时，这些攻击仍然会成功，并且很大一部分攻击者确实表现出一些老练和注重细节的迹象。作为对横向网络钓鱼攻击成功率的估计，我们数据集中至少有 11% 的攻击者成功入侵了至少一个其他员工帐户。就更精细的策略而言，31% 的横向网络钓鱼者会投入一些人力来逃避检测或提高攻击的成功率。此外，我们数据集中超过 80% 的攻击发生在被劫持账户的正常工作时间。综上所述，我们的结果表明，横向网络钓鱼攻击构成了一种普遍的企业威胁，其复杂程度仍有增长空间。

除了探索事件粒度的攻击（如§5），本节还探讨了在研究不同的攻击者行为时以横向钓鱼者（被劫持的帐户）为粒度的攻击。如第节所述2，行业从业者通常将此类被劫持的账户称为 ATO，在本节中，我们将被劫持账户、横向钓鱼者和 ATO 作为同义词使用。

6.1 横向网络钓鱼的规模和成功

规模：我们的数据集包含 154 个被劫持帐户发送的 1,902 封不同的横向网络钓鱼电子邮件。[4] 在我们的数据集中，共有 33 个组织经历过至少一次横向网络钓鱼事件：其中 23 个组织来自对已知横向网络钓鱼事件的一组企业进行抽样（§3），而其余 10 个来自我们从普通人群中抽样的 69 个组织。假设我们的随机样本反映了更广泛的企业群体，超过 14% 的组织在 7 个月的时间跨度内至少经历过一次横向网络钓鱼事件。此外，基于

[4]不同的电子邮件是通过具有完全唯一的元组（发件人、主题、时间戳和收件人）来定义的。

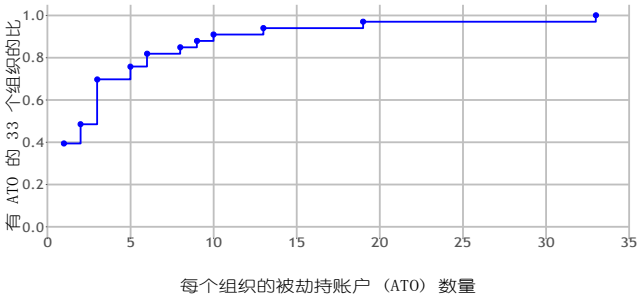

图 4：具有 x 个被劫持帐户且发送至少一封横向网络钓鱼电子邮件的组织的比例。 13 个组织只有 1 个 ATO；其余 20 个发现来自 2 个以上 ATO 的横向网络钓鱼（§6.1).

图上4，在我们的数据集中，超过 60% 的受感染组织经历了来自至少两个被劫持员工帐户的横向网络钓鱼攻击。鉴于我们的攻击集可能包含漏报（因此低估了攻击的普遍性），这些数字表明横向网络钓鱼攻击在企业组织中普遍存在。

成功的攻击：根据我们的数据集，我们不确定攻击是否成功。但是，我们使用以下方法保守地（低估）估计了横向网络钓鱼的成功率。基于这一程序，我们估计至少有 11% 的横向网络钓鱼者成功地破坏了至少一个新的企业帐户。

让 Alice 和 Bob 代表同一组织的两个不同的 ATO，其中 P_A和 P_B分别代表 Alice 和 Bob 的一封钓鱼邮件，ReplyB 代表 Bob 对从 Alice 收到的横向钓鱼邮件的回复。直观地，我们的方法得出结论，如果 (1) Bob 收到来自 Alice 的网络钓鱼电子邮件，(2) 在收到 Alice 的网络钓鱼邮件后不久，Bob 随后发送了他自己的网络钓鱼邮件，以及 (3) 我们有强有力的证据证明这两个员工的网络钓鱼电子邮件是相关的（反映在下面的标准 3 和 4 中）。

形式上，如果以下所有条件都为真，我们就说 P_A成功入侵了 Bob 的账户：

1. Bob 是 P_A的接收者
2. 在收到 P_A后，Bob 随后发送了他自己的横向钓鱼邮件 (P_B)
3. 满足以下两个条件之一：
   (a) P_B和 P_A使用相似的网络钓鱼内容：如果两次攻击使用相同的主题，或者他们使用的两个网络钓鱼 URL 属于同一个完全限定的域
   (b) Bob 向 P_A发送了回复 (ReplyB)，他的回复表明他中了 Alice 的攻击，而 Bob 在他自己的攻击 (P_B) 之前发送了 ReplyB
4. 满足以下两个条件之一：
   (a) P_B在 Bob 收到 P_A后两天内发送
   (b) P_B和 P_A使用相同的网络钓鱼消息或其网络钓鱼 URL 的路径遵循几乎相同的结构（例如，‘ http://X.com/z/office365/index.html ’与 ‘http://Y.com/z/office365/index.html’)

拆开最终标准（#4），在第一种情况（4.a）中，我们根据先前的光照确定了两天的到达间隔阈值特征 [21, 22]，这表明 50% 的用户会在 2 天内回复电子邮件，大约 75% 的点击垃圾邮件的用户会在 2 天内回复。假设网络钓鱼遵循类似的时间常数来确定收件人采取行动所需的时间，2 天代表在 P_A和 P_B之间建立链接的保守阈值。同时，之前的两项工作都表明，存在一长串用户需要数周时间才能阅读电子邮件并对其采取行动。第二部分（4.b）试图通过提高 Alice 和 Bob 的攻击之间的相似性要求来解决这个长尾问题，然后得出前者导致后者的结论。对于被启发式 4.b 标记的成功攻击者，P_A和 P_B之间观察到的最长时间间隔为 17 天，这属于基于上述文献的合理时间尺度。

从这种方法中，我们得出结论，17 个 ATO 成功地破坏了至少 23 个未来的 ATO。虽然我们的程序可能会错误地识别出攻击者同时危害了 Alice 和 Bob（而不是通过 Alice 的帐户危害 Bob 的帐户）的情况，但前两个条件（要求 Bob 是 Alice 的网络钓鱼电子邮件的最近收件人）有助于减少此错误。我们的程序可能低估了横向网络钓鱼攻击的一般成功率，因为它没有识别攻击者随后没有使用 Bob 的帐户发送网络钓鱼电子邮件的成功攻击，也没有考虑我们数据集中的漏报或我们可见范围之外的攻击（例如，外部组织的收件人的妥协）。

6.2 收件人定位

在本节中，我们估计数据集横向网络钓鱼攻击的转换率，并讨论反映数据集中大多数攻击者行为的四种收件人定位策略。

收件人数量和估计的对话率：我们数据集中的横向网络钓鱼者累计联系了 101,276 个唯一收件人，其中 41,740 人与 ATO 属于同一组织。如图5，超过 94% 的攻击者将他们的网络钓鱼电子邮件发送给超过 100 个收件人；对于所有横向网络钓鱼者的总体而言，这个百分比可能高估了高“收件人数量”攻击者的普遍性，因为我们的检测器利用了与收件人相关的特征。

以数百人为目标为攻击者提供了更多的潜在受害者，但同时也带来了风险，即接收者会检测到攻击并将其标记给他们的安全团队或

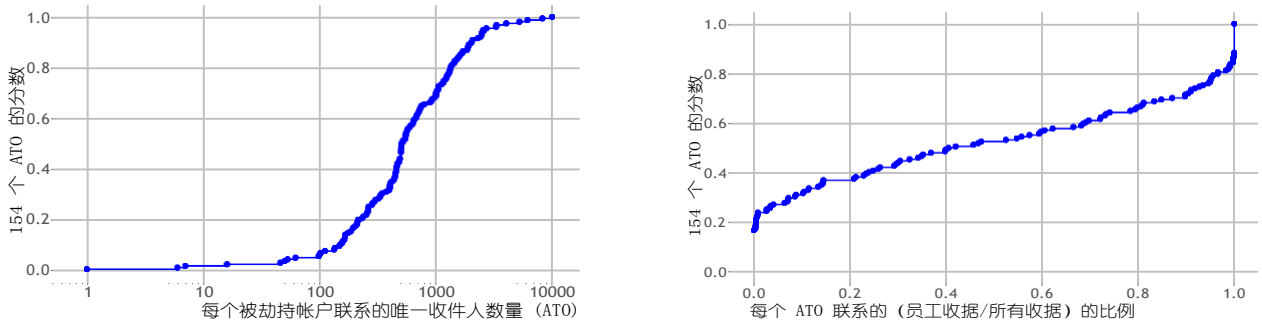

图 5：左侧 CDF 显示每个 ATO 的网络钓鱼收件人总数的分布。右侧的 CDF 显示了 ATO 的比例，其中 x% 的总收件人组由同事组成。

他们的其他收件人（例如，通过回复所有人）。为了隔离他们的受害者并尽量减少其他收件人互相警告的能力，我们发现攻击者经常通过群发密件抄送或通过许多个人电子邮件联系他们的收件人。除了这种遏制策略外，我们还估计我们数据集的横向网络钓鱼攻击很难愚弄单个员工，因此可能需要针对许多收件人来劫持一个新帐户。前面部分6.1，我们发现 17 个 ATO 成功入侵了 23 个新账户。查看他们成功劫持的帐户数量除以他们所针对的同事数量，我们的攻击者的对话率中位数是每 542 名同事一个新劫持的帐户；每次成功妥协，对话率最高的攻击者平均联系 26 名员工。我们警告说，我们确定攻击是否成功的方法（§6.1）并未涵盖所有情况，因此我们的会话率也可能低估了这些攻击在实践中的成功率。但是，如果我们估计的转化率准确地接近真实转化率，就可以解释为什么这些攻击者联系了这么多收件人，尽管被发现的风险增加了。

收件人定位策略：有趣的是，我们知道一些横向网络钓鱼者通过利用被劫持帐户中的信息来定位熟悉的用户来选择他们的一组受害者；例如，将他们的攻击发送到帐户“联系簿”的子集。不幸的是，我们的数据集不包含有关攻击者为选择网络钓鱼收件人而执行的任何侦察操作的信息（例如，明确搜索用户的通讯录或最近的收件人）。

相反，我们凭经验探索数据集攻击者的接收者集，以确定这些攻击者如何选择他们的受害者集的合理策略。表中总结了四种收件人定位策略3（在下面解释），反映了我们数据集中除六个攻击者之外的所有攻击者的行为。为了帮助评估收件人和 ATO 是否存在有意义的关系，我们计算每个 ATO 的最近联系人：ATO 在 ATO 网络钓鱼电子邮件之前的 30 天内至少向其发送过一封电子邮件的所有电子邮件地址的集合。虽然一些攻击者（28.6%）特别

收件人定位策略	# ATO
账户无关	63
全组织范围	39
横向组织	2
目标收件人	44
无定论	6

表 3：每个 ATO 的收件人定位策略摘要（§6.2）.

以帐户的许多最近联系人为目标，大多数横向网络钓鱼者似乎对联系许多任意收件人或向被劫持帐户组织的大部分人发送网络钓鱼电子邮件更感兴趣。

账户无关的攻击者：从目标最少的行为开始，我们数据集中的 63 个 ATO 将他们的攻击发送给范围广泛的收件人，其中大多数人似乎与被劫持的账户没有密切关系。我们称这个群体为 Accountagnostic 攻击者，并使用两种启发式方法识别他们。

首先，如果少于 1% 的收件人与 ATO 属于同一组织，并且进一步探索他们的收件人并没有揭示与该帐户的紧密联系，我们将攻击者归类为与帐户无关的攻击者。检查图中右侧的图表5，37 个 ATO 的目标收件人集合中只有不到 1% 的收件人与 ATO 属于同一组织。为了排除这些攻击者的收件人仍然与帐户相关的可能性，我们计算了出现在每个 ATO 最近联系人中的收件人的比例；对于所有 37 个可能的帐户不可知 ATO，只有不到 17% 的攻击总收件人出现在他们最近的联系人中。在这 37 个与帐户无关的候选 ATO 中，其中 33 个联系 10 个或更多组织的收件人（唯一的收件人电子邮件域），其中 2 个专门针对 Gmail 或 Hotmail 帐户，其余 2 个 ATO 最好被描述为横向组织攻击者（以下）。[5] 剔除2个横向组织攻击者，本次识别出的35个ATO

[5]我们的扩展技术报告提供了所有 ATO 联系的收件人域的分布[17].

第一标准将他们的攻击主要发送给外部收件人，这些收件人属于许多不同的组织或专门用于个人电子邮件托管服务（例如 Gmail 和 Hotmail），并且这些收件人中只有一小部分出现在 ATO 最近的联系人中；因此，我们将这 35 名攻击者标记为与帐户无关。

其次，我们通过搜索 ATO 总收件人中不到 50% 的人也属于 ATO 组织，并且 ATO 与许多不同组织的收件人联系的攻击者来扩大对帐户不可知攻击者的搜索；具体来说，ATO 的网络钓鱼收件人所属的唯一域是 ATO 最近联系人中所有电子邮件地址的两倍多。这个搜索标识——

提交了 63 个 ATO。为了过滤掉这组中可能利用被劫持账户的最近联系人的攻击者，我们排除了任何 ATO，其中超过 17% 的攻击总收件人也出现在 ATO 的最近联系人中（17% 是 ATO 从第一个开始的最大百分比）会计不可知启发式）。应用最后一个条件后，我们的第二个启发式方法识别出 54 个与帐户无关的攻击者。

根据这两个标准对 ATO 进行组合和重复数据删除，得出总共 63 个与帐户无关的攻击者（40.9%）：横向网络钓鱼者主要针对与被劫持帐户或其组织没有密切关系的收件人。

横向组织攻击者：在探索潜在的与帐户无关的 ATO 期间，我们发现了 2 名攻击者，我们将其标记为不同类别：横向组织攻击者。在这两种情况下，只有不到 1% 的攻击者收件人与 ATO 属于同一组织，但每个攻击者的收件人都属于与 ATO 组织属于同一行业的组织。收件人的这一主题特征表明，他们采取了一种深思熟虑的策略，在目标行业的组织中传播，因此，我们将他们归类为横向组织攻击者。

组织范围的攻击者：Office 365 提供了一个“组”功能，列出了一个帐户所属的不同组 [30]。对于某些企业，此功能列举了组织中的大部分（如果不是全部）员工。因此，希望广泛传播网络钓鱼网络的横向网络钓鱼者可能会采用一种简单的策略，即向组织中的每个人发送攻击。我们将这些 ATO 称为组织范围的攻击者，并通过两种方式识别它们。

首先，我们搜索任何攻击者，其中至少一半的钓鱼收件人属于 ATO 的组织，并且该组织的至少 50% 的员工收到了钓鱼邮件（即，钓鱼者的大多数受害者是员工，攻击者将其作为目标企业的大多数）；这次搜索总共产生了 16 个 ATO。我们通过构建一组发送或接收电子邮件的所有员工电子邮件地址来估计组织的员工列表

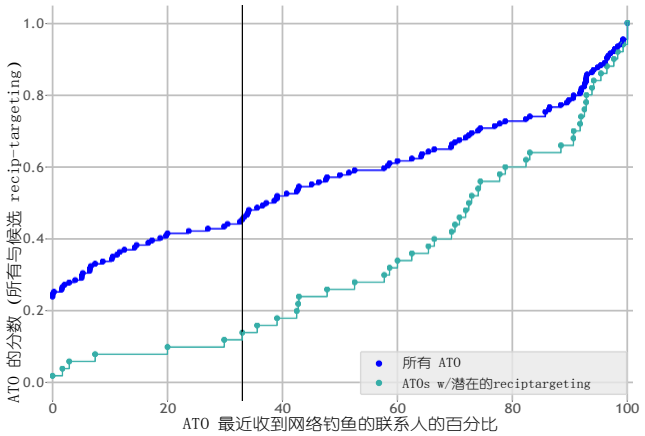

图 6：CDF：x 轴显示 ATO 最近的联系人中有多少百分比收到了横向网络钓鱼电子邮件（§6.2）。底部蓝绿色图表过滤 ATO，以排除任何被识别为会计不可知、横向组织和组织范围攻击者的 ATO；在垂直黑线上，88% 的过滤后的 ATO 将网络钓鱼电子邮件发送到至少 x = 33% 的最近联系人地址。

在网络钓鱼事件的整个月中的任何人。[6] 对于所有这 16 个 ATO，只有不到 11% 的目标收件人也出现在他们最近的联系人中。再加上这些 ATO 中的每一个都与 1,300 多个收件人联系，他们的行为表明他们的最初目标侧重于对尽可能多的企业收件人进行网络钓鱼，而不是针对特别接近被劫持帐户的用户。因此，我们将他们归类为组织范围内的攻击者。

我们的第二个启发式寻找收件人集几乎完全由同事组成的攻击者，但组织的大多数人不一定会收到网络钓鱼电子邮件。重温图5，36 个候选组织范围内的 ATO 将超过 95% 的网络钓鱼电子邮件发送给了同事收件人。但是，我们再次需要排除并说明利用被劫持账户最近联系人的 ATO。从之前讨论的第一个组织范围的启发式中，我们看到该启发式的组织范围内攻击者的接收者中只有不到 11% 来自 ATO 最近的联系人。使用此值作为第二个组织范围内攻击者候选集的最终阈值，我们确定了 29 个组织范围内的攻击者，其中超过 95% 的收件人属于 ATO 的组织，但只有不到 11% 的收件人来自 ATO 最近的联系人；这种组合表明攻击者主要寻求危害其他员工，但这些员工不一定与被劫持的帐户有个人联系。

对上面的两组横向网络钓鱼者进行聚合和重复数据删除，总共产生了 39 个组织范围内的攻击者（25.3%），他们利用被劫持帐户中的信息来攻击许多同事。

[6] 由于服务地址、邮寄列表别名和人员流失，此集合可能高估了实际员工组。

目标收件人攻击者：对于其余未分类的 50 个 ATO，我们无法最终确定攻击的收件人目标策略，因为我们的数据集没有为我们提供攻击者可用的全套信息和操作。尽管如此，图6 提供了一些证据表明，剩下的攻击者中有 44 人确实利用了被劫持帐户的先前关系。具体来说，44 名攻击者向 ATO 最近联系的地址中至少 33% 的地址发送了攻击。[7] 由于这些 ATO 至少向 ATO 最近联系的每 3 个收件人中的 1 个发送了攻击，因此这些攻击者似乎有兴趣将大部分与被劫持账户有已知联系的用户作为目标。因此，我们将这 44 个 ATO 标记为目标接收者攻击者。

6.3 消息内容：定制和主题 由于横向网络钓鱼者控制着合法的员工帐户，这些攻击者可以轻松地挖掘最近的电子邮件以制作个性化的鱼叉式网络钓鱼消息。为了了解攻击者在多大程度上利用了他们的网络钓鱼攻击中的特权访问权限，本节描述了我们在横向网络钓鱼消息中看到的定制级别。总体而言，我们数据集中只有 7% 的事件在其消息中包含目标内容。在使用非目标内容的网络钓鱼电子邮件中，我们数据集中的攻击者依靠两种主要的叙述（欺骗性借口）来引诱受害者执行恶意操作。这两个结果的结合表明，就目前而言，这些攻击者（跨越数十个组织）看到了机会主义网络钓鱼尽可能多的收件人的更多价值，而不是花时间挖掘被劫持的帐户以获得个性化的鱼叉式网络钓鱼饲料。

内容裁剪：在分析我们数据集中的网络钓鱼消息时，我们发现两个维度恰当地描述了不同级别的内容裁剪和定制。第一个维度，“主题定制”，描述了电子邮件的主题或主要思想对受害者或组织的个性化程度。第二个维度，“名称定制”，描述了攻击者如何具体称呼受害者（例如，“亲爱的用户”与“亲爱的鲍勃”）。对于这两个维度中的每一个，我们列举了三个不同级别的定制，并在下面提供了一个匿名消息片段；我们使用 Bob 来指代攻击的接收者之一，并使用 FooCorp 来指代 Bob 工作的公司。

1. 主题剪裁：混乱主题对受害者或组织的独特性和相关性：

[7]在检查和应用帐户不可知攻击者和组织范围攻击者的阈值时，我们使用了一个略有不同的分数：有多少 ATO 的网络钓鱼收件人也出现在他们最近的联系人中？在这里，我们试图捕获专门针对大量熟悉的收件人做出特定努力的攻击者。因此，我们查看 ATO 最近联系人中收到钓鱼邮件的比例，其中分母反映了 ATO 最近联系人中的用户数量，而不是 ATO 的钓鱼邮件收件人总数。

	通用的	企业	有针对性
没有命名	90	35	9
组织命名	23	16	4
收件人姓名		030	

表 4：每个消息定制类别的事件数量分布（§6.3）。这些列对应于消息主题与受害者或组织相关的独特性和具体性。这些行对应于网络钓鱼电子邮件是否明确指定了收件人或组织。

   (a) 通用网络钓鱼主题：可以发送给任何用户的非特定消息（“您有一个新的共享文档可用。”）

   (b) 与企业广泛相关的主题：一条针对企业环境的消息，但如果攻击者在许多其他组织中使用它也有意义（“更新的工作计划。请分发给您的团队。”）

   (c) 目标主题：主题明显依赖于有关收件人或组织的具体细节的消息（“请参阅随附的关于 FooCorp 25 周年的公告。”，其中 FooCorp 已经存在整整 25 年。）

2. 名字剪裁：钓鱼邮件是否特别使用收件人或组织的名字：

   (a) 非个性化命名：攻击没有提到组织或收件人的名字（“亲爱的用户，我们检测到您的邮箱设置有错误……”）

   (b) 具体命名的组织：攻击仅提及组织，但未提及收件人（“来自 FooCorp 的新安全电子邮件...”）

   (c) 收件人特别命名：攻击在电子邮件中特别使用受害者的姓名（“Bob，请查看随附的采购订单...”）

总而言之，该分类法将网络钓鱼内容分为九种不同的定制类别；桌子4 显示我们数据集的 180 起事件中有多少属于每个类别。从这个分类中，出现了两个有趣的观察结果。首先，只有 3 起事件（1.7%）实际提到了收件人的名字。由于我们数据集中的大多数 ATO（94%）向至少 100 个收件人发送电子邮件，因此攻击者需要利用某种形式的自动化来发送数百封单独的电子邮件并自定义每封电子邮件的命名。根据我们的结果，这些攻击者似乎并不认为这是一项值得的投资。例如，他们可能担心发送许多单独的电子邮件可能会触发反垃圾邮件或反网络钓鱼机制，我们在一个试图发送数百封单独电子邮件的 ATO 的案例中观察到这种情况。第二，

单词	# 事件	单词	# 事件
文档	89	发送	44
看法	76	审查	43
附	56	分享	37
点击	55	帐户	36
符号	50	使用权	34

表 5：所有 180 起横向网络钓鱼事件中最常用的 10 个词。

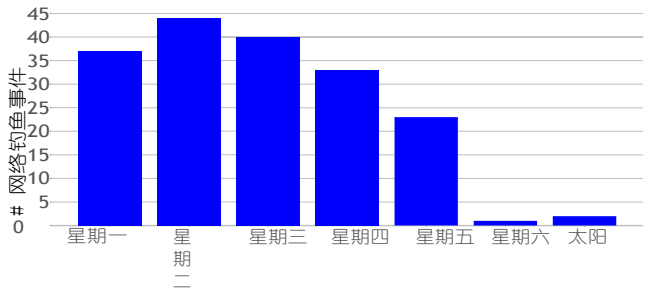

图 7：一周中每天的横向网络钓鱼事件数量。

查看表的最后一列4，只有 13 个事件（7%）在其消息中使用目标内容。绝大多数（92.7%）的事件选择更通用的消息，攻击者可以在大量组织中部署这些消息，只需进行最小的更改（例如，仅更改受害组织的名称）。

虽然我们的攻击数据集捕获了所有横向网络钓鱼攻击的有限视图，但它仍然反映了 7 个月时间范围内 33 个组织中所有已知的横向网络钓鱼事件。因此，尽管数据存在局限性，但我们的结果表明，很大一部分横向网络钓鱼者并未充分利用其受损帐户的资源（即历史电子邮件）来制作个性化的鱼叉式网络钓鱼消息。这一发现表明，这些攻击者的行为更像是投机取巧的网络犯罪分子，而不是顽强的 APT 或民族国家。然而，鉴于安全的军备竞赛和进化性质，这些横向网络钓鱼者将来可能会利用帐户之前的电子邮件来制作更具针对性的内容，从而提高攻击的复杂性和效力。

主题内容（诱饵）：当用一定程度的定制标记每个网络钓鱼事件时，我们注意到我们数据集中的网络钓鱼消息绝大多数依赖于两种欺骗性借口（诱饵）之一：(1)收件人的帐户（并敦促他们点击链接来解决问题）；(2)一条消息，通知接收者有新的/更新的/共享的文档。对于后一种“文档”诱惑，文档的性质和特殊性随内容定制的水平而变化。例如，使用通用主题定制的攻击只会提及模糊的文档，而使用企业相关定制的攻击会将术语切换为发票、采购订单或其他一些通用但与工作相关的文档。

为了进一步描述这种行为，我们计算了数据集中网络钓鱼消息中出现频率最高的词。首先，我们为每个事件选择了一封网络钓鱼电子邮件，以防止具有许多相同电子邮件的事件影响（夸大）其诱饵的流行度。接下来，我们规范化了每封电子邮件的文本：我们删除了自动生成的文本（例如，用户签名），将所有单词小写，删除标点符号，并丢弃所有不常用的英语单词；所有这些都可以使用开源库来完成，例如 Talon [27] 和 NLTK [5]。最后，我们构建了一组在我们的事件中出现在任何网络钓鱼电子邮件中的所有单词，并计算每个单词出现的事件数量。

有趣的是，我们数据集的网络钓鱼消息使用了相对较小的单词库：我们数据集中每封钓鱼邮件的文本中只有 444 个不同的常用英语单词（即，每封网络钓鱼电子邮件的文本都由这组 444 个单词组成）字）。相比之下，我们数据集中的 1,000 封电子邮件的随机样本包含总共 2,516 个不同的词，其中只有 176 封电子邮件完全由网络钓鱼术语集中的词组成。

除了横跨横向网络钓鱼电子邮件的这一小部分总词外，除一个事件外，所有事件都至少包含前 20 个词中的一个，说明了对我们确定的两个主要诱饵的依赖。我们的扩展技术报告显示了每个词的出现分布 [17]。仅关注前十个词和使用它们的事件数（表5），这两个主题诱饵的主导地位变得明显。指示“共享文档”诱饵的词，例如“文档”、“查看”、“附加”和“评论”，每种都出现在超过 23% 的事件中，其中最受欢迎的（文档）出现在近一半的事件中所有事件。同样，我们在前十名中也看到了与帐户相关的诱饵中的许多词：'access'、'sign'（来自'sign on'）和'account'。

总的来说，虽然我们的数据集包含多个有针对性的网络钓鱼消息实例，但我们观察到的大多数横向网络钓鱼电子邮件都依赖于更普通的诱饵，攻击者可以毫不费力地在多个组织中重复使用这些诱饵。我们看到这种行为在数十个不同的组织中反复出现，这一事实表明要么出现了一种新的但可访问的企业网络钓鱼形式，要么是“普通”网络犯罪分子执行网络钓鱼攻击的方式发生了演变（从使用聪明的外部帐户转移欺骗受感染但合法的帐户）。

6.4 横向网络钓鱼的时间方面

由于攻击者可能不在与被劫持帐户相同的地理区域生活或经营，因此先前的工作建议使用捕获网络钓鱼电子邮件中固有的异常时间属性的功能 [11, 15, 41]。与这种直觉相反，在我们的数据集中，大多数横向网络钓鱼攻击发生在一天和一周的“正常”时间。首先，对于 98% 的横向网络钓鱼事件，攻击者在

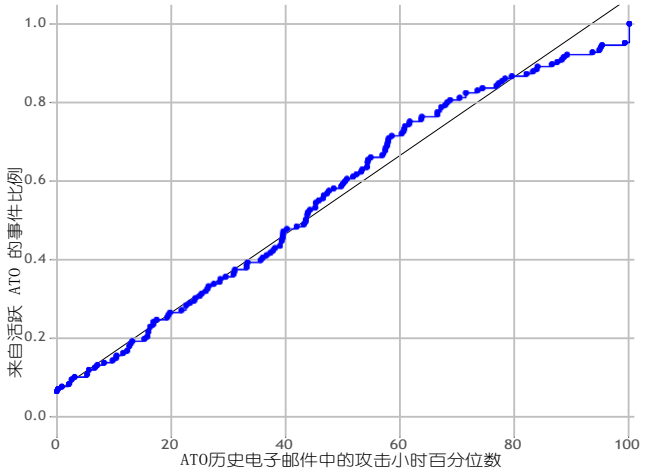

图 8：来自活动 AT 的事件比例的 CDF，其中一天中的时间（小时）在前 30 天内发送 ATO 良性电子邮件的小时数的第 x 个百分位数内。活动 AT 是被劫持的帐户，在横向网络钓鱼电子邮件之前的 30 天内发送了至少 1 封非网络钓鱼电子邮件。

一个工作日。此外，我们数据集中的大多数攻击者在真实帐户的正常工作时间内发送网络钓鱼电子邮件。

星期几：从图7，除三起横向网络钓鱼事件外，所有事件都发生在一个工作日（周一至周五）内。此模式表明，攻击者在员工通常发送良性电子邮件的同一天发送网络钓鱼电子邮件，并且星期几将提供无效或较弱的检测信号。此外，67% 的事件发生在上半周（周一至周三），这表明我们数据集中的横向网络钓鱼者不遵循民间传说策略，即攻击者倾向于在周五发起攻击（希望利用即将到来的周末安全团队操作减少）[37].

一天中的时间（小时）：除了在正常工作周内操作外，大多数攻击者还倾向于在被劫持帐户的典型工作时间发送横向网络钓鱼电子邮件。为了评估攻击发送时间的（异常）异常，对于每个 ATO，我们收集了该帐户在其第一封横向网络钓鱼电子邮件之前的 30 天内发送的所有电子邮件。然后，我们将每封历史（可能是良性的）电子邮件的发送时间映射到 24 小时刻度的一天中的小时，从而形成了每个被劫持帐户通常发送电子邮件的典型一小时的分布。最后，对于每个横向网络钓鱼事件，我们计算了网络钓鱼电子邮件的一小时相对于 ATO 历史电子邮件的一天中小时分布的百分位数。例如，百分位数为 0 或 100 的网络钓鱼事件发送的时间早于或晚于真实帐户所有者在过去 30 天内发送的任何电子邮件。

在活动 ATO 发送的所有横向网络钓鱼事件中，图8 显示网络钓鱼的一小时百分位数 Dent 的第一封电子邮件发生在与被劫持帐户的历史电子邮件相关的位置。在 180 个事件中，有 15 个事件是由“非活动”（静止）ATO发送的，该 ATO 在其横向网络钓鱼电子邮件之前的所有 30 天内发送了零电子邮件;数字8 不包括这些事件。在活动ATO发送的其余165个事件中，有18个事件完全超出了被劫持帐户的历史营业时间，这表明查找用户在非典型时间发送的电子邮件的功能可以帮助检测这些攻击。但是，对于其余 147 个事件，网络钓鱼电子邮件的一小时均匀覆盖了整个百分位范围。如图所示8，网络钓鱼小时的百分位分布与均匀随机分布的 CDF（直线 y = x 线）非常相似;即，网络钓鱼电子邮件的一小时似乎是从真实帐户的历史小时随机抽取的-

的日分布。该结果表明，对于我们数据集中的大多数事件（180 个事件中的 147 个），ATO 发送攻击的时间不会提供重要的信号，因为它们的发送时间反映了真实用户历史电子邮件活动的时间分布。

因此，根据数据集中的攻击，我们发现存在两个与时间相关的弱特征：搜索突然开始发送可疑电子邮件的静态帐户（15 个事件），以及搜索完全在帐户的历史活动时间窗口之外发送的可疑电子邮件（18 个事件）。除了这两个特征以及它们所反映的一小部分网络钓鱼攻击之外，星期几和一天中的时间都没有提供重要的检测信号。

6.5 攻击者的复杂程度

由于我们数据集的大多数横向网络钓鱼者不会挖掘被劫持帐户的邮箱来制作有针对性的消息，因此人们可能会自然而然地得出结论，这些攻击者是懒惰或不成熟的。但是，在本小节中，我们确定了两种需要投入额外时间和手动工作的复杂行为：攻击者不断与攻击的接收者接触以提高攻击的成功率，以及主动“清理”其网络钓鱼活动的痕迹以试图向帐户的合法所有者隐藏其存在的攻击者。与少数攻击者花时间为一组个性化的收件人制作定制的网络钓鱼消息相反，近三分之一（31%）的攻击者至少参与了这两种复杂行为中的一种。

与潜在受害者的互动：收到网络钓鱼邮件后，一些收件人自然会质疑电子邮件的有效性，并发送回复要求更多信息或保证。虽然懒惰的攻击者可能会忽略这些收件人的回复，但 15 个组织的 27 个 AT 通过发送后续消息向受害者保证网络钓鱼电子邮件的合法性，积极与这些潜在受害者接触。例如，在一个组织中，攻击者始终发送简短的后续消息，例如“是的，我已发送给您”或“是的，

你检查过了吗？在其他情况下，攻击者用更详细的诡计回复：例如，“嗨，[鲍勃]，这是一份关于[X]的文档。打开是安全的。您可以通过使用您的电子邮件地址和密码登录来查看它。

为了查找网络钓鱼者主动跟进其攻击的潜在受害者的情况，我们收集了每个横向网络钓鱼电子邮件线程中的所有邮件，并检查攻击者是否曾经收到并回复收件人的回复（查询）。[8] 我们发现，总共有 107 个 AT 至少收到了一个收件人的回复。在这些回复接收攻击者中，27 个 AT （25%） 对其一个或多个收件人的查询发送了欺骗性的后续回复。

隐蔽性：除了与潜在受害者互动之外，攻击者可能会花费手动工作，通过删除其网络钓鱼电子邮件的任何痕迹来向帐户的真正所有者隐藏他们的存在，特别是因为横向网络钓鱼者似乎在被劫持帐户的正常工作时间内运行（§6.4）.为了估计这些 AT 的数量，我们搜索了以下任何电子邮件是否最终进入被劫持帐户的“废纸篓”文件夹，并在发送或接收后 30 秒内被删除：任何网络钓鱼电子邮件、网络钓鱼电子邮件回复或攻击者发送的后续电子邮件。30 秒阈值将隐蔽行为与修正被盗帐户导致的删除区分开来。总共有 16 个组织中的 30 名攻击者参与了这种规避清理行为。

在以交互方式回应有关其攻击的询问的 27 个 AT 中，只有 9 个也表现出这种隐蔽的清理行为。因此，计算两组攻击者的数量，48 个 AT 至少参与了其中一种行为。

参与复杂行为的攻击者中有相当一部分在我们的数据集中创建了更复杂的攻击图景。鉴于这些攻击者确实投入了专门的和（通常）手动的努力来提高攻击的成功率，为什么他们中的许多人（在我们的数据集中超过 90%）使用非有针对性的网络钓鱼内容并针对数十到数百个收件人？这种通用行为的一个合理原因是，他们目前使用的简单方法在他们的经济模式下工作得很好：投入更多的时间来开发更量身定制的网络钓鱼电子邮件并不能提供足够的经济价值。另一个原因可能是横向网络钓鱼攻击的增长反映了网络钓鱼领域的演变，以前“简单”的外部网络钓鱼者已经转向通过横向网络钓鱼发送攻击，因为来自（欺骗）外部帐户的攻击变得太困难了， 由于用户意识和/或针对外部网络钓鱼的更好技术缓解措施。最终，根据我们工作的数据集，我们无法很好地回答为什么这么多横向网络钓鱼者采用简单的攻击，并将其作为一个有趣的问题留给未来的工作来探索。

[8] Office 365 包含一个对话 ID 字段，并且同一线程中的所有电子邮件（原始电子邮件和所有答复）都将分配相同的对话 ID 值。

7 总结

在这项工作中，我们首次对来自92个企业组织的超过1亿封员工发送的电子邮件进行了横向网络钓鱼攻击的大规模表征。我们还开发并评估了一种新的检测器，该检测器发现了许多已知的横向网络钓鱼攻击，以及数十种未报告的攻击，同时产生了少量的误报。通过对数据集中攻击的详细分析，我们发现了许多重要的发现，这些发现为我们的心理模型提供了企业面临的威胁，并为未来的防御指明了方向。我们的研究表明，14% 的随机抽样组织（从小到大）在七个月内经历过横向网络钓鱼攻击，攻击者至少有 11% 的时间成功破坏了新帐户。我们发现并量化了几种主题收件人定位策略和欺骗性内容叙述;虽然一些攻击者进行有针对性的攻击， 大多数遵循采用非个性化网络钓鱼攻击的策略，这些攻击可以在不同组织中轻松使用。尽管在定制和定位攻击方面明显缺乏复杂性，但我们数据集中 31% 的横向网络钓鱼者参与了某种形式的复杂行为，旨在提高他们的成功率或掩盖他们被劫持帐户的真正所有者的存在。此外，相对于合法帐户的历史电子邮件行为，超过 80% 的攻击发生在典型的工作日和小时内;这表明这些攻击者要么与他们劫持的帐户位于相似的时区，要么在他们的 VICS 正常工作时间内协同努力进行操作。最终，我们的工作首次提供了对新兴的、广泛的企业网络钓鱼攻击形式的大规模见解，并阐明了防御这种强大威胁的技术和未来想法。

确认

我们感谢匿名审稿人Itay Bleier和我们的牧羊人Gianluca Stringhini的宝贵反馈。这项工作部分得到了休利特基金会通过长期网络安全中心的支持，NSF资助CNS-1237265和CNS-1705050，NSF GRFP奖学金，Irwin Mark和Joan Klein Jacobs信息与计算机科学主席（UCSD），谷歌和Facebook的慷慨礼物，Facebook奖学金以及UCSD网络系统中心的运营支持。

参考文献

[1] 赛义德·阿布-尼梅、达里奥·纳帕、王欣磊和苏库·奈尔。用于网络钓鱼检测的机器学习技术的比较。2007年第二届ACM电子犯罪程序。

[2] 凯文·阿利克斯、特加文德·比斯扬德、雅克·克莱因和伊夫·勒特拉昂。您的训练数据集是否相关？2015年第7届施普林格ESSoS论文集。

[3] 安德烈·伯格霍尔茨、张郑浩、格哈德·帕斯、弗兰克·赖查茨和西贤·史特博。使用基于模型的功能改进了网络钓鱼检测。2008年第五届CEAS论文集。

[4] 詹姆斯·伯格斯特拉和约书亚·本吉奥。随机搜索超参数优化。JMLR，13（2月），2012.

[5] 史蒂文·伯德、爱德华·洛珀和伊万·克莱因。自然语言工具包。https://www.nltk.org/, 2019.

[6] 埃利·布尔施泰因、博尔巴拉·本科、丹尼尔·马戈利斯、塔德克·彼得拉谢克、安迪·阿切尔、艾伦·阿基诺、安德烈亚斯·皮西尔-利迪斯和斯特凡·萨维奇。手工欺诈和勒索：野外手动帐户劫持。第 14 届 ACM IMC 论文集，2014 年。

[7] 阿萨夫·西顿。威胁聚焦：Office 365 帐户接管 — 新的"内部威胁"。https:blog.barracuda.com/2017/08/30/threat-聚光灯-办公室-365-帐户-妥协-新的内部威胁/，八月 2017.

[8] Asaf Cidon，Lior Gavish，Itay Bleier，Nadia Korshun，Marco Schweighauser和Alexey Tsitkin。高精度检测企业电子邮件泄露。2019年第28届Usenix Security会议记录。

[9] 域名密钥已识别的邮件。https://en.wikipedia.org/wiki/DomainKeys_Identified（维基百科.org/wiki域密钥标识）_邮件. 访问时间：2018-11-01。

[10] Sevtap Duman，Kubra Kalkan-Cakmakci，Manuel Egele，William Robertson和Engin Kirda。电子邮件探查器：具有电子邮件标题和文体功能的鱼叉式网络钓鱼过滤。第 40 届 IEEE COMPSAC 论文集，2016 年。

[11] 曼努埃尔·埃格勒、詹卢卡·斯特林希尼、克里斯托弗·克鲁格尔和乔瓦尼·维尼亚。COMPA：检测社交网络上的受感染帐户。2013年第20届ISOC NDSS论文集。

[12] 联邦调查局。商业电子邮件泄露了 120 亿美元的骗局，2018 年 7 月。https://www.ic3.gov/media/2018/180712.aspx.

[13] 伊恩·费特、诺曼·萨德和安东尼·托马西奇。学习检测网络钓鱼电子邮件。2007年第16届ACM WWW论文集。

[14] Sujata Garera，Niels Provos，Monica Chew和Aviel D Rubin。用于检测和衡量网络钓鱼攻击的框架。2007年第5届ACM WORM论文集。

[15] 雨果·加斯康、斯蒂芬·乌尔里希、本杰明·斯特里特和康拉德·里克。字里行间阅读：内容鱼叉式网络钓鱼电子邮件的不可知检测。第 21 届施普林格突袭的会议记录，2018 年。

[16] 谷歌。分类：大鹏和 AUC。https://developers.google.com/machine- 学习/速成班/分类/ROC-和-AUC, 2019.

[17] Grant Ho、Asaf Cidon、Lior Gavish、Marco Schweighauser、Vern Paxson、Stefan Savage、Geoffrey M. Voelker和David Wagner。大规模检测和表征横向网络钓鱼（扩展报告）。在arxiv，2019年。

[18] Grant Ho、Aashish Sharma、Mobin Javed、Vern Paxson和David Wagner。在企业设置中检测凭据鱼叉式网络钓鱼攻击。第 26 届 USENIX 安全会议记录，2017 年。

[19] 胡璇、李邦怀、张洋、周长玲、马浩.从图形拓扑的角度检测受损的电子邮件帐户。2016年第11届ACM CFI论文集。

[20] 丹·哈伯德。思科伞100万。https:umbrella.cisco.com/blog/2016/12/14/cisco-雨伞-100万/，2016 年 12 月。

[21] 克里斯·卡尼奇、克里斯蒂安·克雷比奇、基里尔·列夫琴科、布兰登·恩莱特、杰弗里·沃尔克、弗恩·帕克森和斯特凡·萨维奇。Spamalytics：垃圾邮件营销转换的实证分析。第15届ACM CCS论文集，2008年。

[22] 托马斯·卡拉吉安尼斯和米兰·沃伊诺维奇。大型企业的电子邮件信息流。技术报告，微软研究院，2008年。

[23] 马哈茂德·孔吉、优素福·伊拉克人和安德鲁·琼斯。缓解鱼叉式网络钓鱼攻击：基于内容的作者身份识别框架。第6届IEEE ICITST论文集，2011年。

[24] FT实验室。清醒的一天。https://labs.ft.com/2013/05/a-清醒日/？MHQ5J=E6，2013 年 5 月。

[25] 史蒂文斯·勒布隆德、塞德里克·吉尔伯特、乌特卡什·乌帕德亚、曼努埃尔·戈麦斯·罗德里格斯和大卫·霍夫内斯。社会工程漏洞利用文档生态系统的广泛视图。第 24 届 ISOC NDSS 论文集，2017 年。

[26] 史蒂文斯·勒布隆德、阿迪娜·乌里特斯克、塞德里克·吉尔伯特、蔡郑亮、普拉泰克·萨克塞纳和恩金·基尔达。通过非政府组织的视角看待有针对性的攻击。第23届USENIX安全会议记录，2014年。

[27] 邮枪队。房子。https://github.com/ 邮件枪/房子，2018.

[28] William R Marczak，John Scott-Railton，Morgan Marquis-Boire和Vern Paxson。当政府攻击对手时：看看参与者和技术。第23届USENIX安全会议记录，2014年。

[29] Microsoft Graph：消息资源类型。https:developer.microsoft.com/en-us/graph/ docs/api-reference/v1.0/resources/message. 访问时间：2018-11-01。

[30] 微软。人员概述 - Outlook Web App。https://support.office.com/en- 我们/文章/人物-概述-展望-网络- app-5fe173cf-e620-4f62-9bf6-da5041f651bf. 访问时间：2018-11-01。

[31] Brad Miller，Alex Kantchelian，Michael Carl Tschantz，Sadia Afroz，Rekha Bachwani，Riyaz Faizullabhoy，Ling Huang，Vaishaal Shankar，Tony Wu，George Yiu，et al. 用于恶意软件检测的审阅者集成和性能测量。2016年第13届施普林格DIMVA论文集。

[32] 耶利米·奥纳奥拉波、恩里科·马里孔蒂和詹卢卡·斯特林希尼。被 pwnd 后会发生什么：了解在野外使用泄露的网络邮件凭据。第 16 届 ACM IMC 论文集，2016 年。

[33] J·帕尔梅。常见的互联网邮件头。https:tools.ietf.org/html/rfc2076.

[34] 费格斯·彭德尔伯里、法比奥·皮拉齐、罗伯托·乔丹尼、约翰内斯·金德和洛伦佐·卡瓦拉罗。Tesseract：消除跨空间和时间的恶意软件分类中的实验偏差。2019年第28届Usenix Security会议记录。

[35] 凯文·波尔森。谷歌破坏了中国对美国高级官员的鱼叉式网络钓鱼攻击。https://www.wired.com/2011/06/gmail-hack/，2011 年 7 月。

[36] 史蒂夫·拉根。Office 365网络钓鱼攻击为此制造了一场持久的内部噩梦。https:www.csoonline.com/article/3225469/office-365-网络钓鱼攻击-创建-持续- 内幕噩梦.html，2017 年 9 月。

[37] 法赫米达·拉希德。不喜欢星期一？攻击者也没有。https://www.csoonline.com/article/3199997/不要喜欢星期一，也不要- 攻击者.html，八月 2017.

[38] 根据新数据重新训练模型。https://docs.aws.amazon.com/machine-learning/latest/dg/根据新数据重新训练模型.html，2019.

[39] 杰夫·约翰·罗伯茨。国土安全部负责人将网络钓鱼列为最大的黑客威胁。http://fortune.com/ 2016/11/20/杰-约翰逊-网络钓鱼/，2016 年 11 月。

[40] 阿帕奇火花。PySpark 决策树分类模型 v2.1.0. http://spark.apache.org/docs/2.1.0/api/python/pyspark。 ML.html？突出显示=功能重要性# Pyspark.ml.分类。 决策树分类模型。功能重要性.

[41] 詹卢卡·斯特林希尼和奥利维尔·托纳德。那不是你：通过行为建模阻止鱼叉式网络钓鱼。第12届施普林格DIMBA论文集，2015年。

[42] Kurt Thomas，Frank Li，Chris Grier和Vern Paxson。连接的后果：描述Twitter上的帐户劫持。第 21 届 ACM CCS 论文集，2014 年。

[43] 丽莎·瓦斯。黑客如何闯入John Podesta，DNC Gmail帐户。https://nakedsecurity.sophos.com/2016/10/25/how-hackers-breakke-in-john-波德斯塔-DNC-Gmail-账户/，2016 年 10 月。

[44] 科林·惠特克、布莱恩·赖纳和玛丽亚·纳齐夫。网络钓鱼页面的大规模自动分类。第17届ISOC NDSS论文集，2010年。

[45] 维基百科。随机森林。https://en.wikipedia.组织/维基/Random_forest，2019.

[46] 金泽特。研究人员发现了隐藏在众目睽睽之下的 rsa 网络钓鱼攻击。https://www.wired.com/2011/ 08/如何被黑客入侵/，2011 年 8 月。

[47] 赵梦辰、安博和克里斯托弗·基金特维尔德。优化个性化电子邮件过滤阈值，以缓解连续的鱼叉式网络钓鱼攻击。2016年第13届AAAI论文集。

A 检测器实施和评估详细信息

A.1 标记网络钓鱼电子邮件

将电子邮件标记为网络钓鱼或良性：手动标记电子邮件时，我们首先检查了五条信息，即电子邮件是否是报告的网络钓鱼事件、邮件内容、标记的可疑 URL，以及其域在上下文中是否有意义、电子邮件的收件人和发件人。除了少数事件外，我们可以通过上述步骤轻松识别网络钓鱼电子邮件。例如：发送给数百个不相关收件人的有关“共享 Office 365 文档”的电子邮件，其文档链接指向 bit.ly 缩短的［非 Microsoft］域;或

描述非 IT 员工发送的“帐户安全问题”的电子邮件，其中“帐户重置”URL 指向不相关的域。对于更困难的情况，我们分析了电子邮件链中的所有回复和转发，如果它收到多个表示警报或可疑的回复/转发，或者如果被劫持的帐户最终发送了回复说他们没有发送网络钓鱼电子邮件，则将其标记为网络钓鱼。最后，如本节所述3.3 我们访问了标记的网络钓鱼电子邮件样本中的非副作用、可疑 URL;我们访问的所有网址都指向插页式警告页面（例如，Google 安全浏览）或欺骗性登录页面。对于我们的检测器标记的电子邮件，但根据检查上述所有信息似乎是良性的，我们保守地将它们标记为误报。在许多情况下，误报是显而易见的;例如，我们的检测器标记的“可疑 URL”出现在发件人签名中并链接到其个人网站的电子邮件。

培训练习与实际网络钓鱼电子邮件：除了区分误报和攻击之外，我们还进行了检查以确保横向网络钓鱼事件代表实际攻击，而不是培训练习。首先，根据横向网络钓鱼电子邮件的标头，我们验证了所有发送帐户都是合法的企业帐户。其次，除了五个攻击帐户外，所有攻击帐户在上个月都发送了一封或多封与网络钓鱼无关的电子邮件。这两点让我们相信网络钓鱼电子邮件来自现有的合法帐户，因此代表了实际攻击;也就是说，培训练习不会劫持现有帐户，因为这可能会导致潜在的声誉损害（我们之前参与的企业安全团队不会这样做）。此外，我们数据集的横向网络钓鱼事件都不是梭子鱼已知的训练演习， 并且没有一个横向网络钓鱼 URL 使用已知安全公司的域。

A.2 模型调优和超参数

大多数机器学习模型（包括随机森林）都要求用户设置控制模型训练过程的各种（超）参数。为了确定分类器的最佳超参数集，我们遵循机器学习最佳实践，对下面列出的超参数的所有组合进行了三重交叉验证网格搜索[4].

1. 树数：50–500，以 50 为步长（即 50、100、150、. . 、450、500）
2. 最大树深：10–100，以 10 为步长
3. 最小叶子大小：1、2、4、8
4. （良性/攻击）电子邮件的下采样率：10、50、100、200

由于我们的训练数据集仅包含几十个事件，因此我们使用三个折叠来确保交叉验证中的每个折叠都包含多个攻击实例。我们的实验使用了一个随机森林模型，该模型有 64 棵树，最大深度为 8，最小叶子大小为 4 个元素，每 1 封攻击电子邮件下采样 200 封良性电子邮件，因为这种配置产生了最高的 AUC 分数 [16].但我们注意到，许多超参数组合产生了类似的结果。