A SURVEY OF PHISHING EMAIL FILTERING TECHNIQUES

> 网络钓鱼邮件过滤技术的调查

1.背景

（1）会议/刊物级别

IEEE Communications Surveys & Tutorials (Volume: 15, Issue: 4, Fourth Quarter 2013)

CCF N

（2）作者团队

Ammar Almomani, B. B. Gupta, Samer Atawneh, A. Meulenberg, and Eman Almomani

（3）论文背景

网络钓鱼电子邮件对电子商务构成了严重的威胁，且呈现上升趋势。过滤钓鱼邮件的方法可以根据攻击流的不同阶段进行分类：网络级保护，认证，客户端工具，用户培训，服务器端过滤和分类等。本文讨论了这些方法的优点和局限性。

H4 网络钓鱼的方式：（网络钓鱼是一种特殊类型的垃圾邮件）

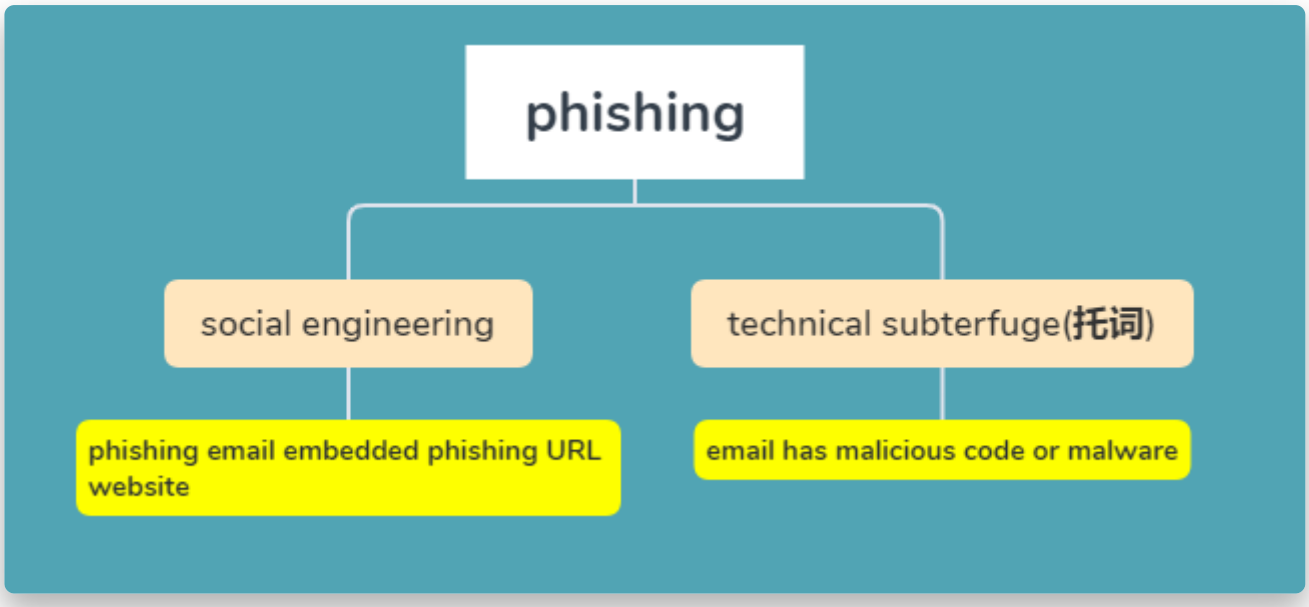
- 欺骗性网络钓鱼

  依赖社会工程学，伪造合法组织的电子邮件，然后再邮件中嵌入虚假链接，诱骗用户访问钓鱼网站。这些虚假网站旨在骗取受害者的财务数据（用户名，密码，信用卡号码和个人信息）

- 基于恶意软件的网络钓鱼

  在用户点击嵌入的URL后， 通过利用用户计算机的漏洞直接获取信息。

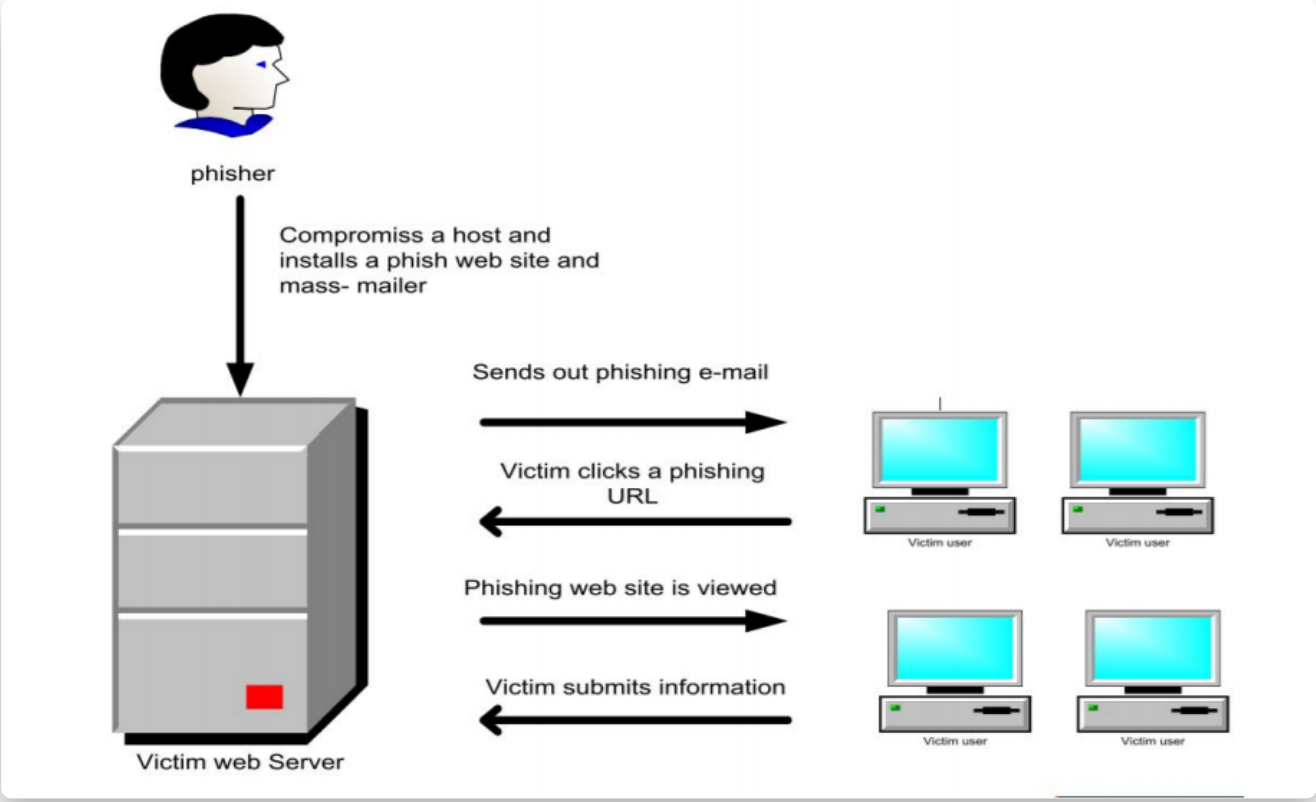
本文聚焦于第一种网络钓鱼方式。



H4 钓鱼邮件的生命周期

网络钓鱼的生命周期通常始于大量发送邮件，试图诱导读者访问邮件中包含的链接。这个阶段的网络钓鱼很像钓鱼。钓鱼者不使用鱼饵和鱼线来钓鱼，而是发送许多电子邮件，希望有一些读者会通过访问电子邮件中包含的链接来“上钩”。通常情况下，电子邮件看起来是合法的，包括一个受欢迎的金融机构的公司标识和合法公司的回邮地址。电子邮件中的链接，“外观”也会在第一眼看上去合法。

来自网络钓鱼网站的网络钓鱼邮件的寿命很短。反网络钓鱼工作组(APWG)收集和存档网络钓鱼攻击的样本。它还关注于识别攻击者的身份，以及由钓鱼攻击、犯罪软件或电子邮件欺骗[10]造成的任何欺诈。



2.钓鱼邮件和特征工程

H3 钓鱼邮件案例

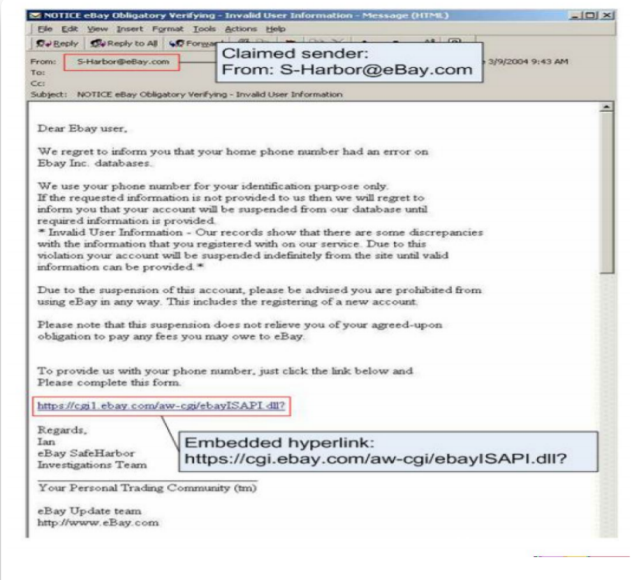


Fig. 3. Screenshot of the eBay phishing email [11]

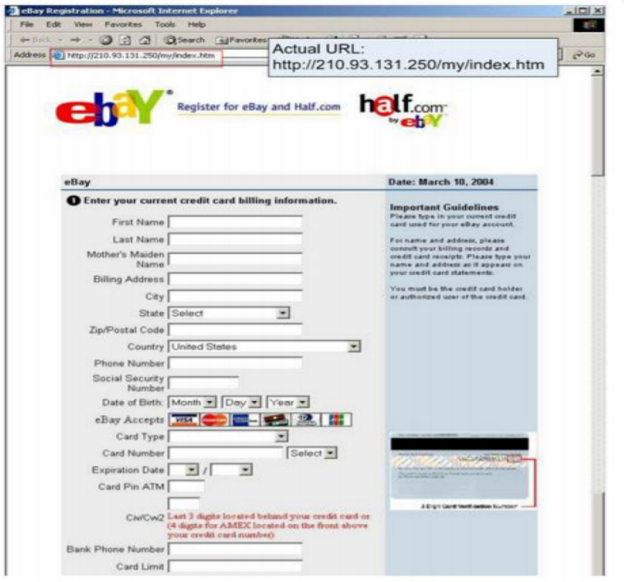


Fig. 4. Screenshot of embedded URL on the eBay phishing web page [11]

钓鱼邮件声称自己来自eBay，表示用户的信息不合法需要修改。该邮件包含一个嵌入URL，似乎是指向合法的eBay网页（这个网页要求用户填写银行卡号码，联系信息，和社会保障号码等）。

H4 发送钓鱼邮件的过程

三个组件：

- 邮件传送代理（MTA）：使用SMTP协议，负责发送和接受电子邮件。
- 邮件投递代理（MDA）：**邮件投递代理”将MTA接收的信件依照信件的流向（送到哪里）将该信件放置到本机账户下的邮件文件中（收件箱），或者再经由MTA将信件送到下个MTA**。
- **MUA（Mail User Agent）**：“邮件用户代理”MUA是用在**客户端的软件**，客户端的计算机无法直接收发邮件，需要通过MUA传递信件，通过各个操作系统提供的MUA才能够使用邮件系统。MUA主要的功能就是接收邮件主机的电子邮件，并提供用户**浏览与编写邮件的功能**，例如outlook，gmail等。

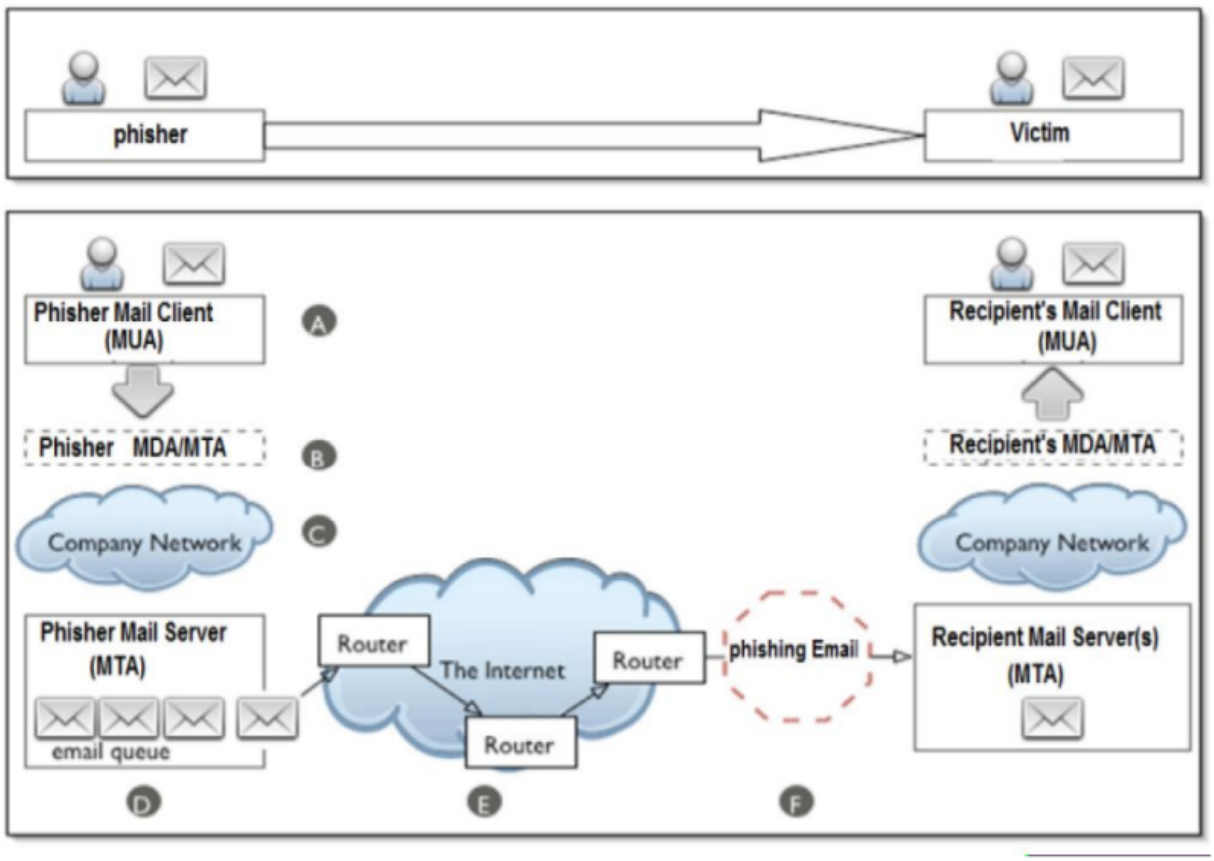


Fig. 5. Phishing email transfer procedure in the computer network

H3 钓鱼邮件分类方法

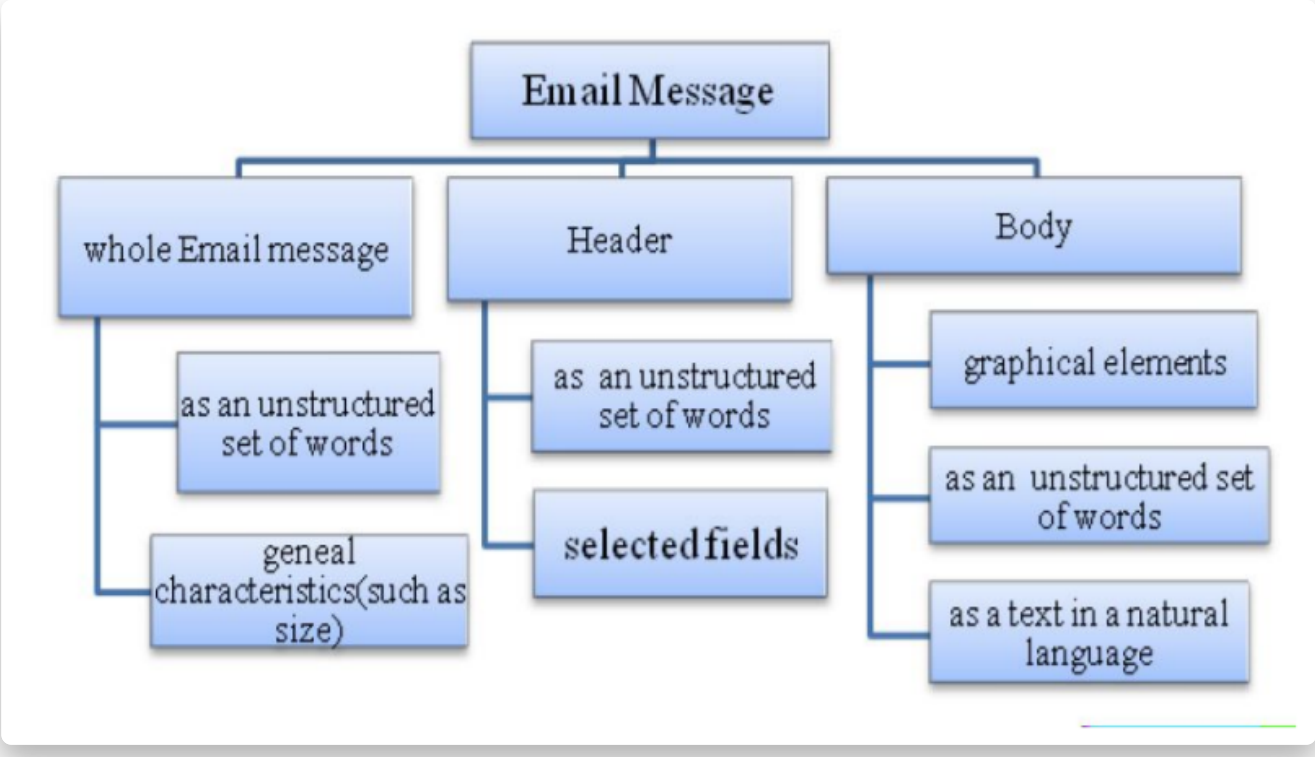
钓鱼邮件过滤器可以定义为一个自动把邮件分为两类（钓鱼邮件和合法邮件）的分类器。

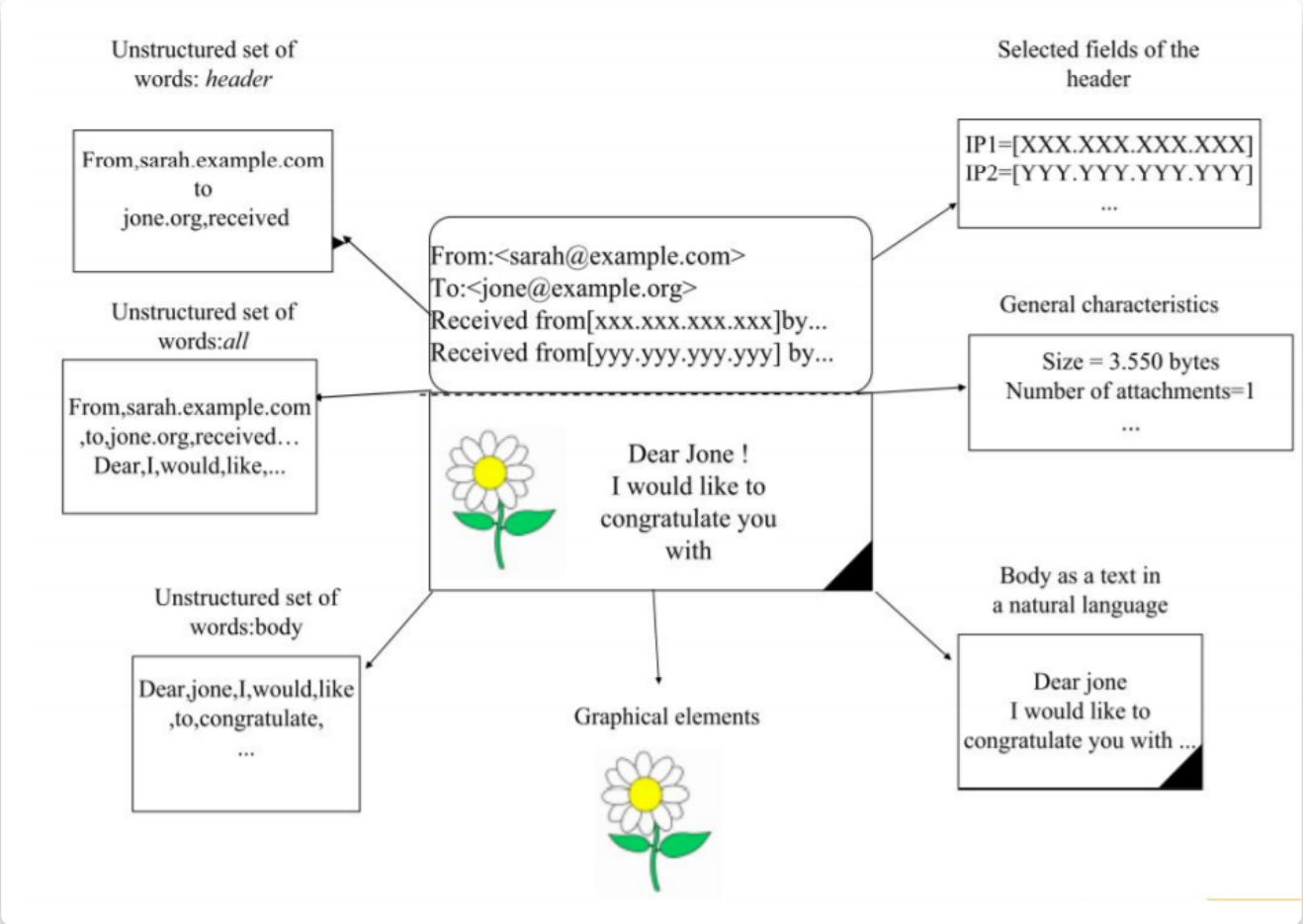
两种过滤方法：

- 检查邮件中是否存在特定单词的关键词过滤
- 使用标记数据集的学习模型

电子邮件的两个组成部分：

- header：由一些结构化的字段组成，例如：From, To, Subject
- body：邮件的实际内容





H3 钓鱼邮件评估方法

N_h：ham email(垃圾邮件)的总数

n_h - H：ham email 被分类为ham的数目

n_h - P：ham email 被分类为phishing的数目

n_p - H：phishing email 被分类为ham的数目

n_p - P：phishing Email 被分类为phishing的数目

评价指标：

- 真正例（TP）：

$$TP = \frac{(n_p -> P)}{N_p}$$

- 真反例（TN）：

$$TN = \frac{(n_h -> H)}{N_h}$$

- 假正例（FP）
- 假反例（FN）

Measure	Formula	Meaning
precision	$= \frac{\|TP\|}{\|TP\| + \|FP\|}$	The percentage of positive predictions that are correct
Recall Sensitivity	$= \frac{\|TP\|}{\|TP\| + \|FN\|}$	The percentage of positive labeled instances that were predicted as positive
Accuracy	$= \frac{\|TP\| + \|TN\|}{\|TP\| + \|TN\| + \|FP\| + \|FN\|}$	The percentage of correct predictions
F-Measure	$= 2 \cdot \frac{precision \cdot recall}{precision + recall}$	A measure of a test's accuracy. It considers both the precision and the recall of the test to compute the score
Total cost ratio	$= \frac{n_P \to H + n_P \to P}{\lambda \cdot n_H \to P + \lambda \cdot n_P \to H}$	λ is the relative cost of the two types of error [13, 17]
Weighted Error (W Err)	$= \frac{\lambda \cdot n_H \to H + n_P \to P}{\lambda \cdot N_H \to N_P}$	According to a specified weight λ [17]

H3 钓鱼邮件特征选取

三种特征：

01. 基本特征：可以不经处理直接从邮件中提取出来的特征
    - 结构特征：从HTML树中提取等

- 链接特征：如IP链接的数量等
- 元素特征：电子邮件中使用的web技术类型：如HTML，脚本，JavaScript
- 垃圾邮件过滤特征
- 单词列表特征：account， update， confirm等

02. 潜在主题模型特征：使用可能在电子邮件中一起出现的词簇（如：在钓鱼邮件中，“点击”和“账户”经常一起出现，而在普通的金融邮件中，“市场”，“计划”和“价格”会在一起出现）
03. 动态马尔科夫链特征：动态马尔可夫链特征是基于词袋模型的文本特征，这些特征为每一类消息的“语言”建模，得出电子邮件归属类别的概率。

H3 特征的筛选和提取

衡量特征有效性并对其排序的算法：

Measure	Formula
Document frequency	$\lvert\{ m_j \lvert m_i \in \mathrm{M} \text{ and } f_i \text{ occurs in } m_j \}\rvert$
Information gain	$\sum_{c \in \{c_{phishingemail}, Cham\}} \sum_{f \in \{fi, \neg fi\}}^{n} \hat{p}(f, c) \log \frac{\hat{p}(f, c)}{\hat{p}(f) \cdot \hat{p}(c)}$

其中M为所有训练消息的集合，Cphishingemail和Cham分别为钓鱼类和垃圾邮件类的标签。fi是一个二元特征(如“点击的词在消息中存在”)，而-fi是特征fi的否定(如“点击的词在消息中不存在”)。所有的概率都是用频率来测量和估计的。

词袋模型主要用于特征提取。在这个模型中，文本（句子或文档）被表示为一个无序的单词集合，而不考虑语法和词序。该模型可以用于表示整个消息或消息的任何部分。

大多数特征选择算法是二分类，一些算法计算同一单词在电子邮件的不同部分出现的次数。

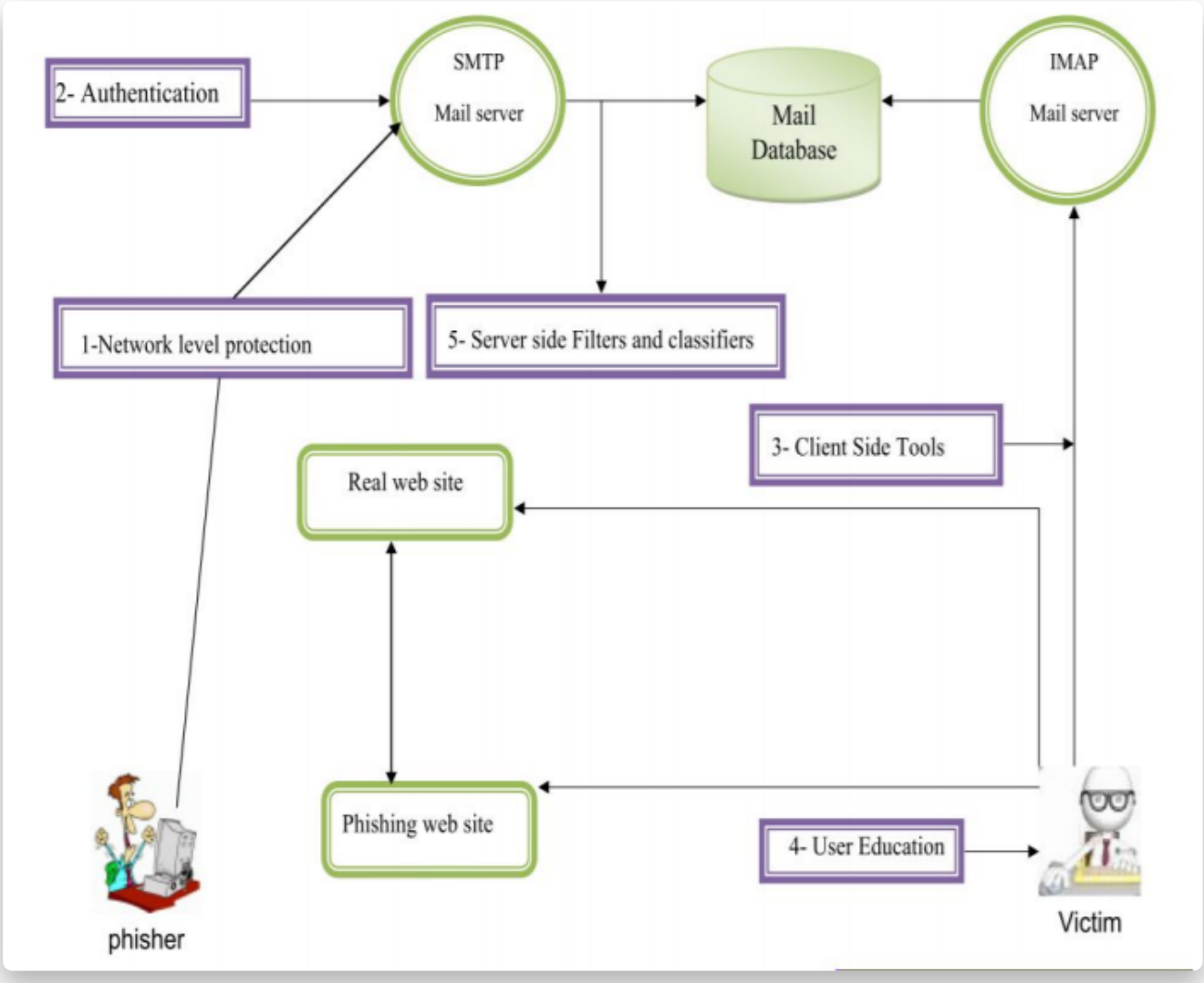
H3 零日钓鱼邮件

> 零日钓鱼邮件是一种新的钓鱼邮件，没有在已有的数据集中出现。

H3 本文与已有工作的比较

		This Work	Related Work	
			Zhang et al. [32]	Zhang et al. [32]
Scope	Network level protection	√		
	Authentication	√		
	Client side tool	√		√
	User education	√		√
	Server side filters and classifiers	√	√	
Surveyed Approaches	Comparison and evaluation	√		√
	Analysis of each techniques	√		
	Attack detection	√	√	
	Vulnerability identification	√	√	√
	Attack protection	√		

3.钓鱼邮件检测方法



如上图所示，根据攻击流中的位置，由五个防御钓鱼的区域。数据流中的其他组件，如简单邮件传输协议(SMTP)和IMAP服务器的协议，是跨Internet协议(IP)传输(电子邮件)的Internet标准。

H3 网络层级别的保护

网络层级别的保护通常是禁止一定范围的IP或域名进入网络。它允许网站管理员阻止来自那些通常发送垃圾邮件或钓鱼邮件的系统的信息。

Tool	Description	Advantages	Disadvantages
Domain name system blacklists [34]	Database used by Internet service providers	An updated list of offending addresses	- Phisher can easily evade this protection technique
Snort [33]	Heuristic/rule engine	Good at detecting level attacks	-Rules require manual adjustments - Does not look at content of message

- “域名黑名单”是互联网服务提供商(ISP)通过研究流量行为来生成和更新的，这种方法本质上是反应性的。攻击者或钓鱼者可以通过控制合法用户的电脑或不断改变IP地址来规避这种保护技术。
- Snort是在网络级别使用的开源软件。Snort中的规则会不断更新，以维护保护力。

H3 身份验证

基于身份验证的方法，旨在确认邮件是否通过有效路径发送，域名是否被钓鱼者欺骗。身份验证在用户和域级别增加了通信的安全性。

两种验证方式：

- 用户级别的身份验证使用密码作为凭据，但是密码验证很容易被攻破
- 域级身份验证在服务端实现

微软引入了一种名为发送者ID的新技术用于域级身份验证(图9)。Yahoo也推出了一种名为域密钥的类似技术。然而，为了使域级身份验证有效，发送方和接收方的提供者必须使用相同的技术。

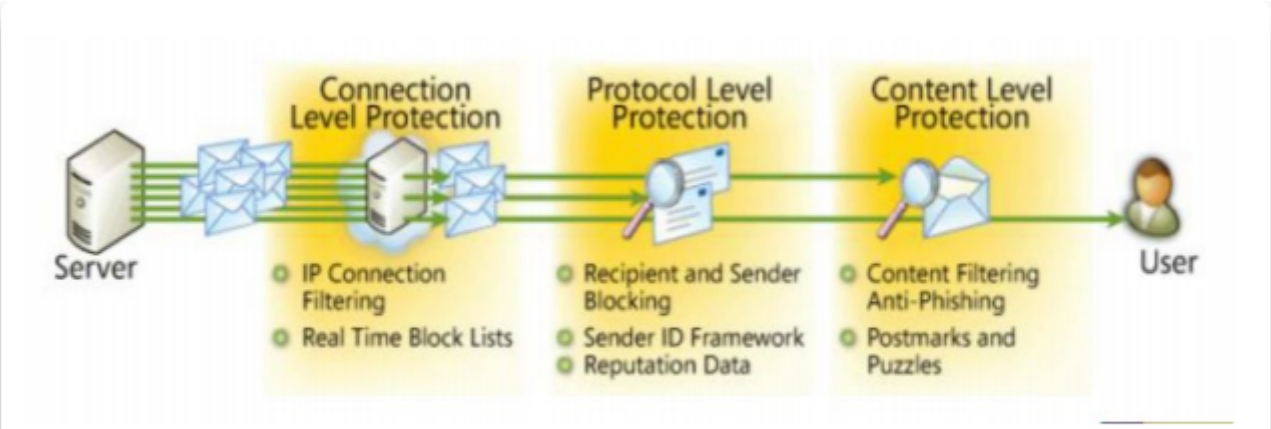


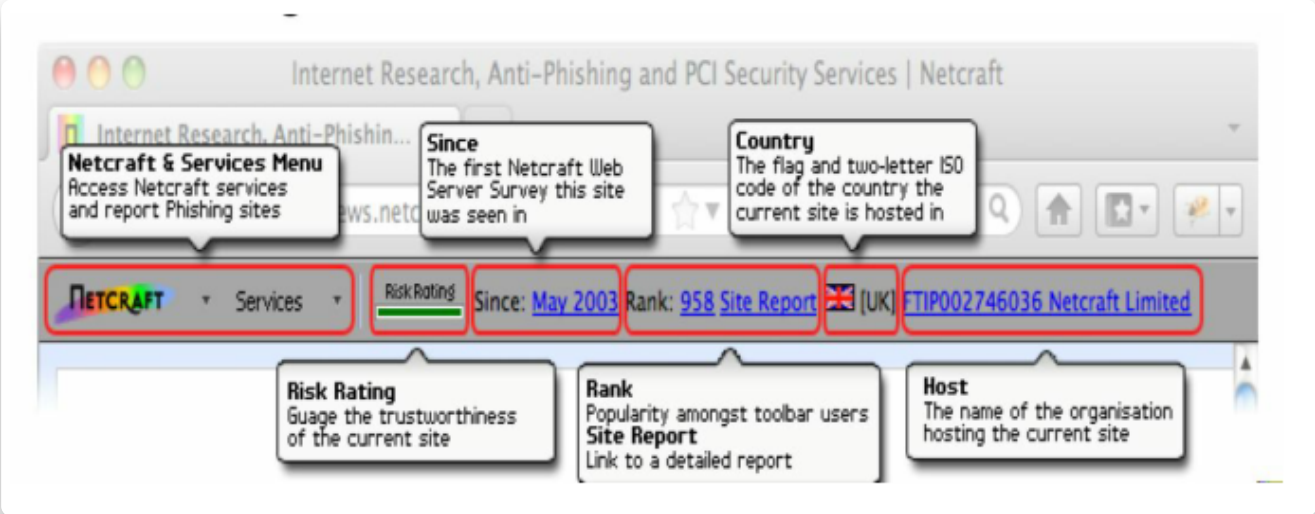
Fig. 9. Microsoft's Sender ID integration into an anti-phishing solution [39]

其他的技术是通过使用数字签名和密码散列来发送带有域名的密码散列。机构将建立一种政策，规定所有与客户的高价值电子邮件通信都要使用授权的私钥进行数字签名。收件人收到电邮后，会使用机构的公开密匙核实电邮的真伪。PGP和S/MIME是数字签名技术的例子。有几位作者推荐使用密钥分发和数字签名来检测钓鱼邮件。但是，目前大多数用户并不使用email认证。

Approach	Strength	Weakness	Used in
Authentication based on user and domain level	-Increases the security of communication	- Both sender and receiver side must employ the same technology	ID [37], Domain Key [38]
Email authentication by digital signature and password hashing	-Less complexity -No need of cooperation bet ween email domains	-At present, most users do not use email authentication	Hotmail, Yahoo, Gmail [43]
Transaction Authentication Numbers (TANs)	-Less complexity -No need of cooperation between email domains	-Affected by man-in-the-middle attacks -Requires substantial infrastructure, time and cost	Bank of Austria [44]

H3 客户端工具

在客户端操作的工具包括用户概要过滤器和基于浏览器的工具栏。其中，SpoofGuard [45]，NetCraft46，CallingID [47]， CloudMark [48]， eBay toolbar[49]， IE phishing filter[50]是其中的一些客户端工具。其中包括对网络钓鱼和直接探测网络钓鱼的“网络浏览器”进行攻击的研究。其他技术还在客户端工具中提出了解决方案，其中包括域检查、URL检查、页面内容、算法和社区输入。这些工具是设计和训练的，使用典型的网络钓鱼网站url的原型，用一个对话框警告用户。



图中清楚地显示了被访问网站的信息，以帮助我们评估欺诈性url(例如，真正的citibank.com或barclays.co.uk网站不太可能托管在前苏联)。通常，这些工具还依赖黑名单和白名单，这是一种通过检查邮件中嵌入的网址或直接检查网站来防止钓鱼攻击的技术。

Approach	Strength	Weakness	Used in
White-listing	Accepts legitimate email only	high false positive	IE, Mozilla Firefox browsers [56, 57]
Black-listing	Good with well-known phishing Web sites	high false negatives	IE, Mozilla Firefox browsers [56, 57]

H3 用户培训

基于社会反应方法的用户教育，一般取决于提高对网络钓鱼攻击的意识和教育水平，特别是对网络钓鱼邮件的意识和教育水平。

Methods	Authority	Attractive	Impressive
Online Material	√	-	-
Online Test	-	√	√
Contextual Training	-	-	√

H3 服务端过滤和分类器

通常，基于基于内容的过滤方法的服务器端过滤器被认为是对抗零日攻击的最佳选择。因此，大多数研究者尝试从这边[21]来解决零日攻击。通常，这种技术依赖于一组提取的钓鱼邮件特征。这些特征通过适应统计分类器在机器学习算法上进行训练，以区分标记为ham(合法)电子邮件或钓鱼电子邮件。之后，可以在电子邮件流上使用这个分类器来预测新收到的电子邮件的类别。

01. 基于词袋模型的方法：
    - 支持向量机（SVM）
    - K近邻算法（KNN）
    - 朴素贝叶斯分类器
    - Boosting算法：决策树
    - 频率-逆文档频率(TF- IDF)
02. 多分类器算法
03. 基于特征的分类器模型:这些方法构建完整的模型，能够使用许多自适应算法和分类器创建新的特征，从而产生最终结果
04. 聚类方法
05. 多层系统
06. 进化的连接主义系统

总结

这篇论文的重点是一个全面的调查钓鱼邮件攻击及其解决方案。目前的保护技术无法阻止钓鱼攻击，特别是“零日攻击”。通过文献调查，总结并评论了各种保护措施，包括:

- 网络层级别的保护
- 身份认证技术
- 客户端工具和过滤器
- 用户教育
- 服务器端过滤和分类