Towards Understanding the Adoption of Anti-Spoofing Protocols in Email Systems

Hang Hu
Virginia Tech
hanghu@vt.edu

Peng Peng
Virginia Tech
pengp17@vt.edu

Gang Wang
Virginia Tech
gangwang@vt.edu

***Abstract*—Email spoofing is a critical step in phishing attacks, where the attacker impersonates someone that the victim knows or trusts. Even today, email providers still face key challenges to detect or prevent spoofing, despite the years of efforts to design and develop anti-spoofing protocols (*e.g.*, SPF, DKIM, DMARC). The key problem is that anti-spoofing protocols are not widely adopted, especially for the new DMARC protocol (5.1%). In this paper, we seek to understand the reasons behind the low adoption rates of anti-spoofing protocols. We conduct a user study with N=9 email administrators from different institutions to understand their perceptions towards anti-spoofing protocols. Our result suggests that email administrators are aware of and concerned about the technical weaknesses in SPF, DKIM, and DMARC that can easily cause errors (*e.g.*, blocking legitimate emails). Email administrators believe the current protocol adoption lacks the crucial mass due to the protocol defects, weak incentives, and practical deployment challenges. Based on these results, we discuss the key implications to protocol designers, email providers and users, and future research directions to mitigate the email spoofing threats.**

I. Introduction

Phishing attack has been a persistent threat to the Internet. Recently, this threat has been significantly escalated due to its heavy involvement in massive data breaches [33], ransomware outbreaks [22], and even political campaigns [9]. For example, spear phishing emails have been used in nearly half of the recent 2000 data breaches, responsible for leaking billions of data records [33].

Email spoofing is a critical step in phishing attacks where the attacker impersonates someone that the victim knows or trusts. By spoofing the email address of a reputable organization or a close friend, the attacker has a better chance to deceive the victim [17]. To prevent spoofing, there has been an active effort since the early 2000 to develop, promote, and deploy anti-spoofing protocols. Protocols such as SPF [19], DKIM [5], and DMARC [20] have become the Internet standards, allowing email receivers to verify the sender's identity.

Despite these efforts, however, sending spoofing emails is still surprisingly easy today. As an example, Figure 1 shows a spoofing email where the sender address is set to the domain of the U.S. Citizenship and Immigration Services (USCIS). We crafted and sent this email to our own account in Yahoo (as the victim), and it successfully reached the inbox without triggering any warnings. This is not a coincident as email spoofing is still widely used in real-world phishing attacks [33], [25], [9].

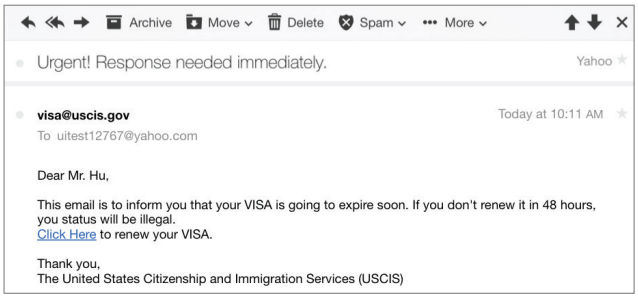

Fig. 1. A spoofing email that impersonates the U.S. Citizenship and Immigration Services (USCIS). We acted as the attacker and sent this email to our own account. The email arrived the inbox without triggering any alert.

The real question is, *why email spoofing is still possible* after years of efforts spent on the defense. In 2015, two measurement studies [8], [12] show that the adoption rates of anti-spoofing protocols are still low. Among Alexa top 1 million domains, only 40% have adopted SPF and only 1% have DMARC. We repeated the same measurement methodology recently in 2018, and found that the adoption rates were not significantly improved (SPF 44.9%, DMARC 5.1%). It is not yet clear what causes the slow progress of adopting anti-spoofing solutions.

In this paper, we seek to understand why anti-spoofing protocols are not widely adopted, particularly from email providers' perspectives. We planned to conduct a user study with email administrators from different institutions, which turned out to be challenging. Part of the reason is that the candidate pool is small. People who can provide insights for our questions need to have extensive experience managing real-world email services. In addition, email administrators often hesitate (or are not allowed) to share details about their anti-phishing/spoofing solutions. To these ends, we send our user study requests to 4000 email administrators of Alexa top domains. We eventually received responses from $N = 9$ administrators from various organizations (universities, payment services, online community websites) who agree to answer open questions either online or through in-person interviews.

Our results show that email administrators are aware of and also concerned about the technical weaknesses of SPF, DKIM and DMARC. Based on interview results and by reading the protocol specifications, we summarize 6 key weaknesses across

the three protocols. These technical weaknesses either allow spoofing emails to bypass the authentication check or block legitimately forwarded emails. The general perception is that these protocols are "helpful", but "cannot solve the spoofing problem completely".

In addition, the email administrators believe that the slow adoption of the protocols is primarily due to the lack of a critical mass. Like many network protocols, the benefits of the anti-spoofing protocols come into existence only if a large number of Internet domains start to adopt the protocols to publish their authentication records. Currently, the incentive of adoption is not strong, especially for Internet domains that don't host emails services (which can still be spoofed).

Finally, the email administrators pointed out the practical challenges to deploy the protocols, particularly, for organizations that use cloud-based email services and large organizations that have many dependent services. Our study participants also shared their thoughts on the possible solutions moving forward. One interesting direction is to improve the current email user interface to support security indicators, and educate users to proactively check email authentication results.

In summary, our work makes three contributions.

- First, we extracted and categorized 6 technical weaknesses in the existing anti-spoofing protocol designs based on our user study (and the protocol specifications). The result provides the taxonomy of the problem.
- Second, through the user study, we provide new insights into the perceived values and concerns of anti-spoofing protocols from email providers' perspectives. These results shed light to the reasons behind the slow adoption of SPF, DKIM, and DMARC, pointing out the directions of improvement moving forward.
- Third, we discuss the key implication of the results to protocol designers, email providers, and users. We discuss the possible solutions at the user-end to make up for the defective server-side authentication.

II. Background and Related Work

In the following, we describe the background of email spoofing attacks and anti-spoofing protocols. Then, we introduce related *technology adoption theories* to set up the contexts for our study.

A. SMTP and Email Spoofing

Simple Mail Transfer Protocol (SMTP) is the Internet standard for email transmission [26], which was designed in 1982. Figure 2 shows a typical email transmission process. A key limitation of SMTP is that it has no built-in security features to prevent people (attackers) from impersonating/spoofing an arbitrary sender address.

To perform a spoofing attack, attackers can manipulate two key fields to send emails. First, after establishing an SMTP connection in step ❷, the attacker can use the "`MAIL FROM`" command and set the sender address to anyone that they want to impersonate. After that, the "`MAIL FROM`" address is inserted into the header as the "Return-Path". In addition, attackers can

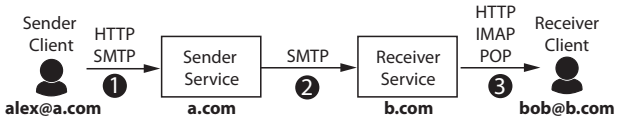

Fig. 2. A typical email transmission process.

modify another field called "`From`" in the email header. This "`From`" field specifies the address that will be displayed on the email interface [28]. When a user receives the email, the user will see the "`From`" address (*e.g.*, *visa@uscis.gov* in Figure 1). If the user replies the email, the reply message will go to the "Return-Path" set by "`MAIL FROM`". Note that the two addresses are not necessarily the same. Email spoofing is a critical step of phishing attacks to gain the victim's trust [27], [15], [17], [6], [7], [11], [14], [29].

B. Anti-Spoofing Protocols

To detect and prevent email spoofing, SMTP extension protocols are proposed including SPF, DKIM and DMARC. All three protocols have been published or standardized by the Internet Engineering Task Force (IETF).

SPF. Sender Policy Framework (SPF) was proposed in early 2000, and standardized in 2014 [19]. SPF allows a domain to publish a list of IPs that are authorized to send emails on its behalf. For instance, the domain `a.com` can publish its SPF record in the DNS. When the receiving server receives the `MAIL FROM` command claiming to be `alex@a.com`, the receiving server can check if the sender IP is listed in the SPF record of `a.com`.

DKIM. DomainKeys Identified Mail (DKIM) was first drafted in 2004 and standardized in 2011 [5]. DKIM uses a public-key based approach to authenticate the email sender and check the email integrity. More specifically, the sender's email service will place a digital signature in the email header signed by the private key associated with the sender's domain. The receiving service can retrieve the sender's public key from DNS to verify the signature. To retrieve a DKIM public key from DNS, one will need the selector information (an attribute in the DKIM signature beside the domain name. The DKIM signature contains the signing algorithm, the signing domain, selector for the DKIM DNS record, signed parts of the email and the actual signature. By verifying the DKIM signature, the receiver can detect if the signed message has been modified, to ensure integrity and authenticity. After DKIM was proposed, there has been research efforts seeking to improve the email authentication procedure [13], [24].

DMARC. Domain-based Message Authentication, Reporting and Conformance (DMARC) was drafted in 2011 and published in 2015 [20]. DMARC is not a standalone protocol but needs to work with SPF and/or DKIM. DMARC allows the domain owner to publish a "failing policy" which specifies what actions the receiver should take when the incoming email fails the DMARC checks. In addition, DMARC requires *identifier alignment* from SPF or DKIM. For SPF, alignment

TABLE I

USER STUDY PARTICIPANTS: 9 EMAIL ADMINISTRATORS. U8 REQUESTED TO CONCEAL THE INSTITUTION TYPE, AND THUS WE KEEP IT AS "ANONYMOUS". FOR EACH OF THEIR EMAIL SERVICES, WE ALSO MEASURED WHETHER THE EMAIL DOMAIN PUBLISHED THE DNS AUTHENTICATION RECORDS (AS THE SENDER) AND WHETHER THE DOMAIN AUTHENTICATE INCOMING EMAILS (AS THE RECEIVER). "✓" MEANS THE MAIL SERVER HAS ADOPTED SPF/DKIM/DMARC. "✗" MEANS THE MAIL SERVER DID HAS NOT ADOPTED SPF/DKIM/DMARC. "/" MEANS NOT APPLICABLE. NOTE THAT WE COULD NOT OBTAIN A MAIL SERVER'S DKIM RECORD FROM THE DNS SINCE THE SELECTOR INFORMATION IS NOT PUBLIC.

UserID	User Study Method	Email Service Type	As Sender: Publish Records?			As Receiver: Authenticate?		
			SPF	DKIM	DMARC	SPF	DKIM	DMARC
U1	In-person Interview	University1 (campus-level)	✓	/	✓	✓	✓	✓
U2	In-person Interview	University1 (department-level)	✗	/	✗	✓	✓	✓
U3	Open-question Survey	Payment System	✓	/	✓	✓	✓	✓
U4	Open-question Survey	Website Hosting Service	✓	/	✗	✗	✓	✗
U5	Open-question Survey	Advertisement Service1	✓	/	✓	✓	✓	✓
U6	Open-question Survey	Advertisement Service2	✓	/	✗	✓	✓	✓
U7	Open-question Survey	University2 (campus-level)	✗	/	✗	✓	✓	✓
U8	Open-question Survey	Anonymous	/	/	/	/	/	/
U9	Open-question Survey	Online Community	✓	/	✗	✓	✓	✓

means that `MAIL FROM` address used for the SPF check should be consistent with the `From` field in the header. For DKIM, alignment means that the domain name in the DKIM signature should match the `From` field. Alignment ensures the email address that user sees matches with the authenticated address.

C. The Low Adoption Rates of Anti-spoofing Protocols

In 2015, two measurement studies have shown that anti-spoofing protocols were not widely used among Internet domains [8], [12]. Among Alexa top 1 million domains [2], only 40% of the domains have published an SPF record and 1% have a DMARC record. DKIM is also not widely adopted based on Gmail's internal estimation [8].

In January 2018, we conducted our own measurements to examine the recent adoption rates for SPF and DMARC, following the same methodology of [12], we find that among Alexa top 1 million domains, 44.9% of the domains have a valid SPF record and 5.1% of the domains have a valid DMARC record. Among the 1 million domains, 79% are Email domains with MX records. We find that 54.3% of the MX domains have a valid SPF record, and 6.0% of the MX domains have a valid DMARC record [16]. Compared with the study conducted in 2015, the adoption rates have increased, but only mildly. Our measurement result raises serious concerns about the effectiveness of the current spoofing defense. We are motivated to further explore the reasons behind the low adoption rates of anti-spoofing protocols.

III. USER STUDY METHODOLOGY

In this paper, we conduct an exploratory study to understand the adoption of anti-spoofing protocols. We qualitatively look into the perceptions of email administrators towards existing anti-spoofing protocols. We primarily focus on two aspects: the perceived usefulness (PU) and the perceived ease-of-use (PEOU), which are the two most important factors for general technology adoption [32], [31], [21]. Below, we introduce the methodology of our user study.

The biggest challenge of our user study is to recruit participants. We need to recruit participants who have real-world experience of operating an email service and/or deploying anti-spoofing protocols. This narrows down the candidate pool to a small and highly specialized user population. In addition, real-world email administrators are often reluctant to share due to the sensitivity of the topic. For many companies and organizations, details about their phishing/spoofing detection systems are non-disclosable.

To address these challenges, we sent our user study requests to a large number of email administrators. More specifically, we contacted the email administrators of Alexa top 4000 domains. In the user study request, we ask about their preferred ways of participation (*e.g.*, survey, phone interviews) and the level of details they feel comfortable to share. In total, we recruit $N = 9$ email providers from different organizations. 7 participants agree to fill in a survey with "open questions" and 2 participants agree to do an in-person interview. In Table I, we list the 9 email administrators and the *type* of their institutions and organizations. Note that `U8` requested to conceal the institution-specific information, and thus we keep it as "anonymous". This small-scale but in-depth user study seeks to provide useful qualitative results and new insights from *protocol users*' perspectives.

To provide the context for each email service that the participant manages, we also performed a quick measurement as shown in Table I. We measured whether the email domain published the corresponding authentication records in DNS (as the sender) and whether the domain performed authentication checks on the incoming emails (as the receiver). As mentioned in II-C, we cannot measure whether an email domain has published the DKIM public key without knowing its selector (marked with "/"). We observe that most of the email services perform all three authentication checks on incoming emails (7 out of 8) and one email service checks DKIM only. However, when acting as the sender domain, only 3 email services published both SPF and DMARC records to the DNS.

For the interview and survey participants, we use the same list of open questions. The difference is that we can ask follow-up questions to the interview participants, but not the survey participants. At the high-level, the open questions fall into the following themes. First, we ask the participants to comment on the email spoofing problem and how they usually

TABLE II
TECHNICAL WEAKNESSES OF SPF, DKIM AND DMARC.

Protocol	Weakness	Problem Description
SPF	P1. Alignment	The SPF verified sender address can be different from the one displayed to users.
	P2. Mail forward	A forwarded email by default cannot pass the SPF test.
	P3. Mailing list	Emails sent to a mailing list by default cannot pass the SPF test.
DKIM	P4. Alignment	The sender domain that signed DKIM can be different from the one user sees.
	P5. Mailing list	Mailing lists often modify the email content, which will fail the DKIM test.
DMARC+SPF	P2. Mail forward	A forwarded email by default cannot pass the SPF test, and thus fails DMARC.
	P3. Mailing list	Emails sent to a mailing list cannot pass SPF and DMARC at the same time.
DMARC+DKIM	P5. Mailing list	Mailing lists often modify the email content, which will fail the DKIM test.
DMARC+SPF+DKIM	P5. Mailing list	SPF always fails; DIKM will fail if the mailing list modifies email content.

detect spoofing attempts. Second, we ask the participants to comment on the value and potential weaknesses of SPF, DKIM and DMARC. Third, we ask about their personal perceptions towards the under-adoption of anti-spoofing protocols and the possible reasons. Fourth, we ask the participants to comment on the possible solutions moving forward to the email spoofing problem.

The survey participants answer the open questions using an online survey website that we set up. The interview participants then have a face-to-face interview session for 45 to 60 minutes. Our study is approved by IRB. We ensure that all the data are properly anonymized and securely stored.

IV. USER STUDY RESULTS

In the following, we discuss our user study results regarding the values and concerns of SPF, DKIM and DMARC, and the possible reasons behind their slow adoption. We group the results into 6 high-level topics.

A. Technical Defects of the Protocols

Email administrators have acknowledged the values of adoption these protocols. However, the most discussed topics are still the technical flaws in SPF, DKIM and DMARC. In the following, we categorize and summarize 5 key weaknesses of the anti-spoofing protocols based on the user study results, as shown in Table II. We have validated these weaknesses by (1) reading and the protocol specifications, and (2) deploying SPF, DKIM and DMARC on our own mail server and running proof-of-concept experiments.

Identifier Alignment (P1, P4). SPF and DKIM both have the problem of "identifier alignment". It means that the sender email address that user sees can be different from the address that is actually used to do perform authentication. Figure 3 shows an example for SPF. For SPF, the authentication focuses on the "Return-Path" and examines whether the sender's IP is listed in the "Return-Path" domain's SPF record. An attacker can set the "Return-Path" domain to her own domain and set her SPF record to pass the authentication. However, what the receiving user sees on the email interface is set by the "From" field. Since SPF does not require the two domains to be the same, then the spoofing email can pass the SPF check while displaying the impersonated address to users. DKIM has a similar problem given that the domain to sign the email with the DKIM key can be different from the domain on the "Return-Path" . DMARC helps to revolve the problem by enforcing the alignment of the identifiers.

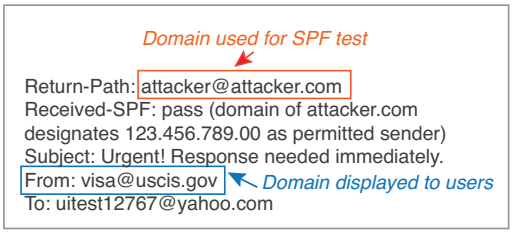

Fig. 3. SPF: SPF test is based on the domain of "Return-Path", which can be different from the domain that the user sees (the "From" field).

Mail Forwarding (P2). Mail forwarding is a problem for SPF. Mail forwarding means one email service automatically forwards emails to another email service. A common scenario is that university students often configure their university email service to forward all their emails to Outlook or Gmail. During Mail forwarding, the email metadata (*e.g.*, "Return-Path") remains unchanged. SPF will fail after mail forwarding because the forwarder's IP will not match the original sender's SPF record. DMARC cannot solve the mail forwarding problem of SPF.

Mailing List (P3, P5). Mailing list is a major problem for both SPF and DKIM. When a message is sent to a mailing list, the mailing list will "broadcast" the message to all the subscribers. This is a similar process as mail forwarding. During this process, the mailing list's IP will become the sender IP, which is different from the original sender's IP. This will lead to SPF failure.

Mailing lists will cause trouble for DKIM because most mailing lists modify the email content before broadcasting it to the subscribers. The common modification is to add a "footer" with the name of the mailing list and a link for un-subscription. Tampering the email content will cause DKIM failure.

DMARC cannot solve the mailing list problem. For mailing lists, DMARC+SPF will be sure to fail: if the "Return-Path" is modified, DMARC will fail due to the misalignment of identifiers; if the "Return-Path" is unmodified, SPF will fail due to the IP mismatch. For DMARC+DKIM, it will fail if the mailing list still has to modify the email content.

In particular, U7pointed out the problem of DKIM beyond just the mailing list problem. U7 stated that DKIM was too sensitive to "benign" changes to the email content such as line rewrapping and URL expansion. These operations that are very common in email services (sometimes for usability purposes), but can easily lead to invalid signatures. The sensitivity of DKIM also discourages email administrators from deploying DMARC (which need to work with DKIM).

> *"U7: DKIM is inherently flawed because semantically meaningless changes to a message can render the signature invalid. For example, the relaxed body canonicalization algorithm is sensitive to line rewrapping, which will invalidate the signature without changing the semantic content of the message. Flaws like this make DKIM signatures fragile, reducing the utility of DKIM and thus lessening the priority of its deployment."*

> *"U7: The fragility of DKIM also affects the utility of DMARC, and thus reducing the priority of its deployment as well."*

B. A Lack of Critical Mass

Email administrators mentioned that there had not been a global consensus that SPF, DKIM or DMARC should be the ultimate solution to stop spoofing. Part of the reason is these protocols are struggling to support common email scenarios such as mail forwarding. Due to the technique weaknesses, the general perception is that SFP, DKIM and DMARC are "helpful" but "cannot solve the spoofing problem completely". *U2* mentioned that potential adopters could be are waiting to see whether enough people would eventually get on board.

> *"U2: It is not the final answer that the industry picked up yet. I felt at this point that enough people haven't really adopted it, it's not worth for me to set it up."*

This reflects a typical bootstrapping challenge, where a "critical mass" is needed in order to facilitate a self-sustaining adoption process [23]. A related notion is the *Network Externalities* (or net effect) [18], [4]. Network externalities mean that an individual adopter can add the value for other people to adopt the same technology. In other words, when more users adopt the same protocol, the value of the protocol to each user will also increase [30]. For anti-spoofing protocols, if more domains publish their SPF/DKIM/DMARC records, it makes easier for other email providers to detect spoofing emails.

C. Benefits Not Significantly Overweight Costs

Email administrators then discussed the deeper reasons for the lack of critical mass. *U1* pointed out that the protocol adopter does not directly benefit from publishing their SPF, DKIM or DMARC records in the DNS. Instead, these DNS records mainly help *other email services* to verify incoming emails and protect the customers (users) of other email services. Domains that publish the DNS records receive the benefit of a better reputation, which is a relatively vague benefit, particularly for domains that don't host email services.

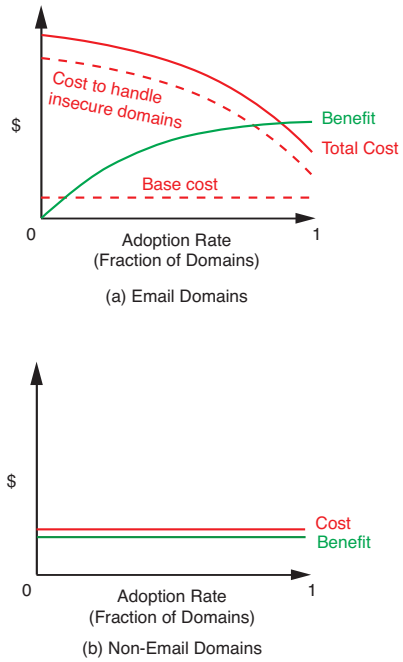

Fig. 4. The adoption model for anti-spoofing protocols. For email domains, the cost and benefit changes as more domains adopt the protocol. For non-email domains, the cost and benefit stay constant.

> *"U1: If I am an email provider, I am not motivated to set up SPF, I am motivated to make sure people who have sent (emails) to my customers have set SPF. I am motivated to evaluate it."*

For popular online services (*e.g.*, social networks, banks), however, they are likely to be motivated to publish SPF, DKIM, and DMARC records to prevent being spoofed and maintain their good reputation (*U2, U3*).

To help to illustrate this challenge, we plot Figure 4, which is a modified version of the Ozment-Schechter model [23]. Ozment-Schechter model depicts the general challenge for network protocols to receive a wide adoption. The model argues that only when the *benefits* to individual adopters overweight the adoption *costs* will the protocol be widely accepted. For network protocols, the per-user benefits may grow as more users adopt the protocol (net effect) [1]. The costs can be either constant or changing (mostly decreasing) as more users adopt the protocol. We have adapted this model to the email spoofing scenarios and created a separate plot for non-email domains (Figure 4(b)).

For email domains (Figure 4(a)), when more domains publish their SPF, DKIM or DMARC records, the benefits for each adopter will increase because more incoming emails can be authenticated. Regarding the costs, there will be a constant *base cost* for deploying the protocol. On top of that, early adopters also need to handle the insecure domains that have not adopted the protocol and those with misconfigurations. With more domains adopting those protocols, there will be fewer emails coming from insecure domains and the cost of

insecure domains will drop. However, this cost cannot reach zero due to the technical issues in these protocols as discussed before.

Figure 4(b) shows a bigger challenge to motivate non-email domains to publish the SPF/DMARC record. For non-email domains (*e.g.*, `office.com`), the benefit of publishing the SPF/DMARC record is to prevent attackers from impersonating the non-email domain and helps the non-email domain to maintain a good reputation. The domain administrators publish the SPF/DMARC records to be a good Internet "citizen" and help other email services to detect spoofing emails. However, these benefits are considered indirect and thus relatively weaker (*U5, U6*). Overall, the cost and benefit model is not in favor of creating a "critical mass" for a wide adoption. The bootstrapping phase is challenging without external enforcement or incentives.

D. Deployment Difficulties in Practice

Even if an email administrator decided to deploy the protocol, there would be other challenges in the way. We summarize the participants' responses from three aspects: (1) a lack of control on the DNS or even the mail servers, (2) the large number of dependency services, (3) a lack of understanding of the protocol and the deployment difficulties.

First, certain services do not have a control over their DNS record. Publishing SPF/DKIM/DMARC record will incur additional overhead to coordinate with their DNS providers (*U1, U4, U9*). In addition, many companies and organizations even don't maintain their own mail servers but rely on cloud-based email services. Using cloud-based email services is convenient without the need the handle challenging tasks such as spam filtering. The drawback is that the organization need to rely on the cloud email service to deploy the anti-spoofing protocols.

> *"U1: So we have very limited control over our DNS. Right now, it is just the difficulty of setting up that DNS."*

Another challenge is that the strict enforcement of certain email protocols requires significant efforts for coordination in big institutions. An email system has many dependent services (*e.g.*, marketing tools) distributed in different departments in a big institution. Deploying a new email protocol requires a non-trivial collaboration effort from different departments.

> *"U7: Strict enforcement requires identifying all the legitimate sources of email using a return address domain. Large, decentralized organizations (e.g. many large universities), will often have organizational units which acquire third-party services involving email, like email marketing tools, without telling central IT. Figuring all this out and putting policies and procedures in place to prevent it is more work than many admins have time for."*

Finally, the participants mentioned that there had been a lack of deep understanding of the anti-spoofing protocols, especially the new protocols such as DMARC. It is difficult to estimate how much effort is needed to deploy and maintain the protocol in practice. *U3* particularly mentioned that there is a general perception that deploying anti-spoofing protocols is difficult. Regardless the actual level of the difficulty, the perceived difficulty makes email administrators hesitated to try (*U3, U9*).

> *"U3: Many people believe that DKIM is hard, and thus don't prioritize deploying it ... Many people don't understand DMARC, how easy it is to deploy, and how effective it is."*

E. Risks of Breaking the Existing System

Participants have discussed the concerns of breaking the existing email system due to unfamiliarity to the protocol. This is particularly true for DMARC (published in 2015). Email providers need to go through careful testing to make sure the protocol does not block legitimate incoming emails, and their own emails are not blocked by others.

> *"U2: Probably because it (DMARC) is still in a testing phase and (people) want to see if it is going to work for them. Relatively it (DMARC) is still pretty new for big businesses and such."*
>
> *"U5: Domains may fear that they've forgotten something and their email may be rejected due to a mistake on their part. "*

These concerns also explain why most protocol adopters (as the sender domain) configure a relaxed SPF/DMARC policy [8], [12] — even if the authentication failed, email providers can still allow email delivery. *U5* expressed that it was quite often for senders to have misconfigurations. It is easier to not enforce the strict policy than to ask the senders to fix their configurations.

> *"U5: Spam filters are relied upon too heavily and it's sometimes easier to pull email from the spam folder than ask someone to fix their SPF record and re-send the email."*

F. Solutions Moving Forward

We asked the participants to comment on the possible solutions moving forward. Most of the email administrators believed that automated detection systems (*e.g.*, anti-spoofing protocols, spam filters, virus scanners) were necessary, but could not fully prevent spoofing or phishing. *U1*, *U2*, *U7*, *U8* and *U9* all have mentioned the importance of user education to raise the awareness of spoofing, and training users to check the email authenticity themselves.

> *"U7: There is no one single way. Technological defenses like content filtering of incoming mail (i.e. spam and virus filtering), are necessary but not sufficient. There is also a need for rigorous training combined with periodic self-phishing (e.g. phishme.com), to raise awareness and identify people who need further training or correction."*
>
> *"U8: User education is the most important way to protect them. I always ask our users to look for the email that seems suspicious and bring it to*

my attention. That way we can prevent malicious intention at earliest possible."

Finally, *U5* expressed the need to have security indicators on the email client. The security indicators are icons or visual cues that are widely used on web browsers to indicate the validity of SSL certificate of websites. A similar email spoofing indicator can be deployed to warn users of emails with unverified sender addresses. In addition, security indicators can also help to highlight the address misalignment of the `Return-Path` and `Mail From` fields for emails that bypassed the SPF check.

"U5: Add the ability for email clients to warn users similar to the way browsers do when users are either presented with a valid extended SSL cert or no SSL cert at all. May also display the from & reply to addresses making it harder to get around SPF record checking."

V. Discussion

So far, we have explored the challenges for SPF, DKIM and DMARC to receive a wide adoption. Next, we discuss the key implications to protocol designers, email providers, and the end users.

A. Implications for Protocol Designers and Promoters

Improving the Perceived Usefulness. The security and usability issues in SPF, DKIM and DMARC negatively impact their perceived usefulness. To improve the perceived usefulness, addressing these security and usability issues becomes the first priority. Currently, an IETF group is working on a new protocol called Authenticated Received Chain (ARC) [3] which is expected to address email forwarding problem and the mailing list problem. However, this also adds to the number of protocols that domain owners need to deploy. New protocols will have their own challenges to be accepted. For example, the DMARC protocol, even though incrementally deployable, only achieved a 4.6% adoption rate in the past two years. A useful protocol will still face the challenge to be widely adopted.

Building the Critical Mass. Currently, there is a lack of strong consensus to deploy anti-spoofing protocols. Like many networking protocols, anti-spoofing protocols will provide key benefits only after enough domains start to publish their SPF, DKIM or DMARC records. To bootstrap the adoption and establish a critical mass, external incentive mechanisms are needed. In theory, we can adjust the rewarding function to provide more benefits to early adopters to create a positive net effect [23]. One possible direction is to learn from the promotion of "HTTPS" among websites [10]: modern browsers will display a trusted icon for websites with valid TLS certificates. Similar security indicators can be added to emails with verified sender domains (by SPF, DKIM and DMARC), to incentive domains to publish the corresponding DNS records. In addition, policymakers or major email providers may also consider enforcing certain sensitive domains (*e.g.*, banks, government agencies) to publish their SPF/DKIM/DMARC records to prevent being impersonated. The challenge is how to realize these ideas without disrupting any of the normal operations of the existing email services.

Reducing the Deployment Difficulty. One direction to improve the adoption rate of anti-spoofing protocols is to make it easy to deploy and configure. Our user study reveals two key problems to address. First, more organizations start to use cloud-based email services (*e.g.*, Google G-Suite, Amazon WorkMail, Office 365). Anti-spoofing protocols should be more cloud-friendly for organizations that don't have full controls on their mail servers. Second, the deployment process should be further simplified and providers email administrators with more controls. The biggest concern from email administrators is that anti-spoofing protocols may reject legitimate emails or get their own emails rejected. One direction of improvement is to allow the protocol to run in a *testing mode* (*e.g.*, in DMARC), allowing email administrators to fully assess the impact before real deployment.

B. Implications for Email Providers

In the short term, email providers are still unlikely to be able to authenticate *all* the incoming emails. While email providers should act as "good Internet citizens" by publishing their own authentication records, it is also necessary to help to "educate" their users to watch out for spoofing emails. Given the current adoption rate of anti-spoofing protocols (and the relaxed protocol configurations), it is likely that email providers will still have to deliver certain unverified emails to the user inbox. Email providers should act more responsibly by providing the authentication results available for the user to check, or proactively warn users of emails that they are not able to verify. Large email providers such as Gmail and Outlook are already moving towards this direction. Currently, Gmail's authentication results are available through the webmail interface, but unfortunately not yet available on the mobile app interface. Further research is needed to improve the current mobile email UI to better support security features.

C. Implications for Users

Given the current situation, users are at the most vulnerable position. Particularly, considering the usability flaws of the existing anti-spoofing protocols, an email that passed the SPF/DKIM checks can still be a spoofed email (*e.g.*, with misaligned addresses). Similarly, emails that failed the SPF/DKIM checks are not necessarily malicious (*e.g.*, forwarded email). To this end, unless the user is fully aware of the authentication details, it is safer for general email users to avoid establishing the trust based on the sender domains. The trustworthiness of the email should be assessed as a whole. It is more reliable to leverage the context of the email exchange, and the external confirmation channels (*e.g.*, calling the sender on the phone) to identify phishing attempts and securely handle critical emails.

VI. Limitations

The scale of the user study is still small, which limits us from producing any statistically significant results. We argue that our contribution is to provide a "qualitative" understanding

of the problem space, which lays the groundwork for future quantitative research. For example, one future direction is to conduct surveys to understand what types of domains are more likely to adopt anti-spoofing protocols, and how domain attributes (*e.g.*, service type, popularity, sensitivity) affect the domain owners' decision.

VII. Conclusion

In this paper, we examine why email spoofing is (still) possible in today's email system. First, our measurement results confirm that anti-spoofing protocols (SPF, DKIM, DMARC) are not widely accepted. Then we qualitatively study the possible reasons for the low adoption rates. By analyzing the discussion threads in IETF and performing user studies with email administrators, we provide a deeper understanding of the perceived value and limitations of anti-spoofing protocols. Our results show that key security and usability limitations are rooted in the protocol design which hurts the perceived usefulness of these protocols. This also makes it difficult to establish a "critical mass" to facilitate a positive net effect for a wider adoption. Moving forward, extensive efforts are needed to address the technical issues in the protocol design and develop external enforcement (or incentives) to bootstrap the protocol adoption. In addition, improved user interfaces are needed for email systems to allow users to proactively check the email authentication results.

Acknowledgments

We would like to thank the anonymous reviewers for their helpful feedback. This project was supported in part by NSF grants CNS-1750101 and CNS-1717028. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of any funding agencies.

References

[1] B. Aboba and D. Thaler, "What makes for a successful protocol?" *RFC5218*, 2008, https://tools.ietf.org/html/rfc5218.

[2] Alexa, 2017, http://www.alexa.com.

[3] K. Andersen, B. Long, S. Jones, and M. Kucherawy, "Authenticated received chain (arc) protocol," ser. Internet-Draft'17, 2017, https://tools.ietf.org/html/draft-ietf-dmarc-arc-protocol-09.

[4] R. Böhme, "Internet protocol adoption: Learning from bitcoin," in *Proc of IAB Workshop on Internet Technology Adoption and Transition (ITAT'13)*, 2013.

[5] D. Crocker, T. Hansen, and M. Kucherawy, "Domainkeys identified mail (dkim) signatures," ser. RFC6376, 2011.

[6] P. Dewan, A. Kashyap, and P. Kumaraguru, "Analyzing social and stylometric features to identify spear phishing emails," in *Proc. of eCrime'14*, 2014.

[7] S. Duman, K. Kalkan-Cakmakci, M. Egele, W. K. Robertson, and E. Kirda, "Emailprofiler: Spearphishing filtering with header and stylometric features of emails," in *Proc. of ACSAC'16*, 2016.

[8] Z. Durumeric, D. Adrian, A. Mirian, J. Kasten, E. Bursztein, N. Lidzborski, K. Thomas, V. Eranti, M. Bailey, and J. A. Halderman, "Neither snow nor rain nor mitm: An empirical analysis of email delivery security," in *Proc. of IMC'15*, 2015.

[9] FBI, "Grizzly steppe: Russian malicious cyber activity," FBI and DHS report, 2016, https://www.us-cert.gov/sites/default/files/publications/JAR_16-20296A_GRIZZLY%20STEPPE-2016-1229.pdf.

[10] A. P. Felt, R. Barnes, A. King, C. Palmer, C. Bentzel, and P. Tabriz, "Measuring HTTPS adoption on the web," in *Proc. of USENIX Security'17*, 2017.

[11] I. Fette, N. Sadeh, and A. Tomasic, "Learning to detect phishing emails," in *Proc. of WWW'07*, 2007.

[12] I. D. Foster, J. Larson, M. Masich, A. C. Snoeren, S. Savage, and K. Levchenko, "Security by any other name: On the effectiveness of provider based email security," in *Proc. of CCS'15*, 2015.

[13] M. T. Goodrich, R. Tamassia, and D. Yao, "Accredited domainkeys: A service architecture for improved email validation," in *Proc. of CEAS'05*, 2005.

[14] G. Ho, A. Sharma, M. Javed, V. Paxson, and D. Wagner, "Detecting credential spearphishing in enterprise settings," in *Proc. of USENIX Security'17*, 2017.

[15] J. Hong, "The state of phishing attacks," *Communications of the ACM*, vol. 55, no. 1, 2012.

[16] H. Hu and G. Wang, "End-to-end measurements of email spoofing attacks," in *Proc. of USENIX Security'18*, 2018.

[17] T. N. Jagatic, N. A. Johnson, M. Jakobsson, and F. Menczer, "Social phishing," *Communications of the ACM*, vol. 50, no. 10, 2007.

[18] M. L. Katz and C. Shapiro, "Technology adoption in the presence of network externalities," *Journal of political economy*, vol. 94, no. 4, pp. 822–841, 1986.

[19] S. Kitterman, "Sender policy framework (spf)," ser. RFC7208, 2014, https://tools.ietf.org/html/rfc7208.

[20] M. Kucherawy and E. Zwicky, "Domain-based message authentication, reporting, and conformance (dmarc)," ser. RFC7489, 2015, https://tools.ietf.org/html/rfc7489.

[21] P. Lai, "The literature review of technology adoption models and theories for the novelty technology," *Journal of Information Systems and Technology Management*, vol. 14, pp. 21 – 38, 04 2017.

[22] Malwarebytes, "Understanding the depth of the global ransomware problem," 2016, https://www.malwarebytes.com/pdf/white-papers/UnderstandingTheDepthOfRansomwareIntheUS.pdf.

[23] A. Ozment and S. E. Schechter, "Bootstrapping the adoption of internet security protocols," in *Proc. of WEIS'06*, 2006.

[24] V. Pathak, D. Yao, and L. Iftode, "Improving email trustworthiness through social-group key authentication," in *Proc. of CEAS'08*, 2008.

[25] PhishMe, "Ransomware delivered by 97% of phishing emails by end of q3 2016 supporting booming cybercrime industry," 2016, https://phishme.com/ransomware-delivered-97-phishing-emails-end-q3-2016-supporting-booming-cybercrime-industry/.

[26] J. B. Postel, "Simple mail transfer protocol," ser. RFC821, 1982, https://tools.ietf.org/html/rfc821.

[27] Proofpoint, "Threat summary and year in review," 2016, https://www.proofpoint.com/sites/default/files/proofpoint_q4_threat_report-final.pdf.

[28] P. Resnick, "Internet message format," ser. RFC5321, 2001, https://www.ietf.org/rfc/rfc2822.txt.

[29] M. Risher, "Protecting you against phishing," Google Security Blog, 2017, https://security.googleblog.com/2017/05/protecting-you-against-phishing.html.

[30] E. M. Rogers, *Diffusion of Innovations, 5th Edition*, 5th ed., 2003.

[31] V. Venkatesh and H. Bala, "Technology acceptance model 3 and a research agenda on interventions," *Decision Sciences*, vol. 39, no. 2, pp. 273–315, 2008.

[32] V. Venkatesh and F. D. Davis, "A theoretical extension of the technology acceptance model: Four longitudinal field studies," *Management Science*, vol. 46, no. 2, pp. 186–204, 2000.

[33] Verizon, "Data breach investigations report," 2017, http://www.verizonenterprise.com/verizon-insights-lab/dbir/2017/.