# Neither Snow Nor Rain Nor MITM... An Empirical Analysis of Mail Delivery Security

Zakir Durumeric[1] David Adrian[1] Ariana Mirian[1] James Kasten[1]
Elie Bursztein[2] Nicolas Lidzborski[2] Kurt Thomas[2]
Vijay Eranti[2] Michael Bailey[3] J. Alex Halderman[1]

[1] University of Michigan [2] Google, Inc. [3] University of Illinois, Urbana Champaign

{zakir, davadria, amirian, jdkasten, jhalderm}@umich.edu
{elieb, nlidz, kurtthomas, vijaye}@google.com
mdbailey@illinois.edu

## ABSTRACT

The SMTP protocol is responsible for carrying some of users' most intimate communication, but like other Internet protocols, authentication and confidentiality were added only as an afterthought. In this work, we present the first report on global adoption rates of SMTP security extensions including STARTTLS, SPF, DKIM, and DMARC. We present data from two perspectives: SMTP server configurations for the Alexa Top Million domains, and over a year of SMTP connections to and from Gmail. We find that the top mail providers (e.g., Gmail, Yahoo, Outlook) all proactively encrypt and authenticate messages. However, these best practices have yet to reach widespread adoption in a long tail of over 700,000 SMTP servers, of which only 35% successfully configure encryption and 1.1% specify a DMARC authentication policy. This security patchwork—paired with SMTP policies that favor failing open to allow gradual deployment—exposes users to attackers who downgrade TLS connections in favor of cleartext and who falsify MX records to reroute messages. We present evidence of such attacks in the wild, highlighting seven countries where more than 20% of inbound Gmail messages arrive in cleartext due to network attackers.

## Keywords

SMTP, Email, Mail, TLS, STARTTLS, DKIM, SPF, DMARC

## 1. INTRODUCTION

Electronic mail carries some of a user's most sensitive communication, including private correspondence, financial details, and password recovery confirmations that can be used to gain access to other critical resources. Users assume that messages are confidential and unforgeable. However, as originally conceived, SMTP, the protocol responsible for relaying messages between mail servers, does not authenticate senders or encrypt mail in transit. Instead, servers support these features through protocol extensions such as STARTTLS, SPF, DKIM, and DMARC. Consequently, adoption of these mechanisms has been gradual, and servers must tolerate both protected and unprotected messages.

In this work, we measure the global adoption of SMTP security features from two perspectives: SMTP connection logs from January 2014 to April 2015 for Gmail, one of the world's largest mail providers; and SMTP server configurations aggregated from the Alexa Top Million domains. At the conclusion of our study, 80% of outgoing and 54% of incoming Gmail messages were protected by TLS. This represents an 84% increase in inbound message security and a 54% increase in outbound message security (strictly for Gmail) over the last year. We find high adoption rates are fueled by a small fraction of popular web mail providers, while adoption lags for over 700,000 SMTP servers associated with the Alexa Top Million. Only 82% of Alexa domains with SMTP servers support TLS and only 34.8% of these are properly configured to allow server authentication. Low adoption stems in part from two of the three most popular SMTP software platforms failing to protect messages with TLS by default. Further, *none* perform certificate validation.

This security patchwork—paired with SMTP policies that favor failing open and transmitting messages in cleartext to allow incremental adoption—enables two techniques for network attackers to intercept mail. In the first attack, network appliances corrupt STARTTLS connection attempts and downgrade messages to non-encrypted channels. In the second attack, DNS servers provide falsified MX records for the SMTP servers of common mail providers. By performing an Internet-wide scan for SMTP servers, we find 41,405 SMTP servers in 4,714 ASes and 193 countries that cannot protect mail from passive eavesdroppers due to STARTTLS corruption. We analyzed the mail sent to Gmail from these hosts and find that in seven countries, more than 20% of all messages are actively prevented from being encrypted. In the most severe case, 96% of messages sent from Tunisia to Gmail were downgraded to cleartext.

To measure another possible attack, we searched the IPv4 address space for DNS servers that provide false addresses for the SMTP servers of five common mail providers. We find 14.6 K publicly accessible DNS servers in 521 ASes and 69 countries providing such false answers. We investigate the messages that Gmail received from these hosts and find that in 193 countries more than 0.01% of messages from each country are transited through these imposter hosts. In the largest case, 0.08% of messages from Slovakia were relayed from a falsified IP, which could have intercepted or altered their contents.

We also study the deployment of three technologies intended to authenticate senders and guard against message spoofing: SPF, DKIM, and DMARC. We find that, during April 2015, 94.4%

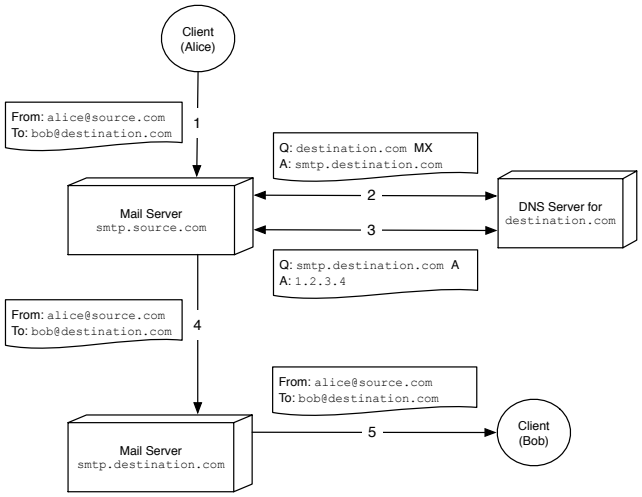


Figure 1: **SMTP Protocol** — Mail is relayed between domains using SMTP. The sending client connects to an outgoing SMTP server and sends mail. The outgoing server performs an MX lookup for the destination domain to identify its incoming mail server and forwards the message.

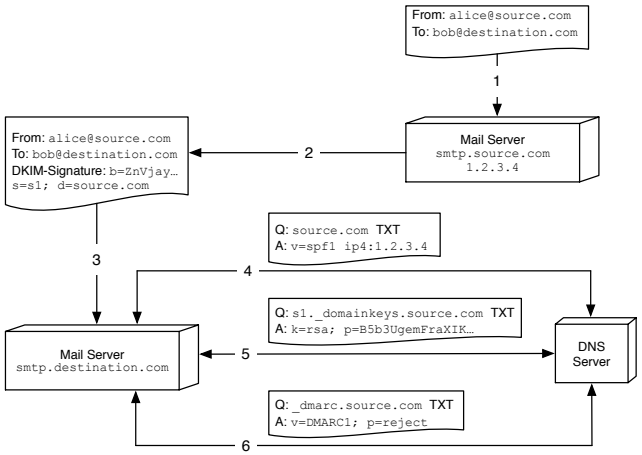


Figure 2: **SPF, DKIM, and DMARC** are used to provide source authentication. The receiving mail server performs an SPF lookup (*4*) to check if the outgoing server is whitelisted, a DKIM lookup (*5*) to determine the public key used in the signature, and a DMARC lookup (*6*) to determine the policy should SPF or DKIM validation fail.

of incoming Gmail messages were protected by a combination of DKIM (82.99%) and SPF (94%). However, among scanned domains, 47% deployed SPF policies and 1.0% deployed DMARC policies. Over 29% of these SPF policies specified more than $2^{16}$ addresses, and 27% include IP ranges that belong to shared cloud providers, which potentially allows collocated VMs to spoof mail.

Drawing on our measurements, we discuss various attack scenarios and challenges, present current proposals for securing mail transport, and propose directions for future research. We hope that our findings can both motivate and inform further work to improve the state of mail security.

## 2. BACKGROUND

Simple Mail Transfer Protocol (SMTP) is the Internet standard for sending and relaying electronic mail [24, 30]. In a simplified scenario—outlined in Figure 1—clients send outgoing mail to their organization's local SMTP server (❶). The local SMTP server performs a DNS lookup for the mail exchange (`MX`) record of the *destination.com* domain, which contains the hostname of the destination mail server, in this case *smtp.destination.com* (❷). The sender's mail server then performs a second DNS lookup for the server's IP address (❸), establishes a connection, and relays the message (❹). The recipient can later retrieve the message using a secondary protocol such as POP3 or IMAP (❺). In practice, mail forwarding, mailing lists, and other scenarios result in messages traversing multiple SMTP relays before being delivered to their final destination.

As when originally conceived in 1981, SMTP lacks support for protecting the confidentiality of messages in transit, and authenticating messages upon receipt. It is vulnerable to both passive observers and active attackers. A passive observer can read message content on the wire while an active attacker can additionally alter messages. Since SMTP's inception, several extensions have been introduced to both protect mail in transit and allow recipients to authenticate the mail they receive.

### 2.1 Protecting Messages in Transit

STARTTLS is an SMTP extension introduced in 2002 that encapsulates SMTP within a TLS session [20]. In a typical STARTTLS session, a client first negotiates a SMTP connection with a server, after which the client sends a `STARTTLS` command that initiates a standard TLS handshake. Mail content, attachments, and any associated metadata are transmitted over this protected channel.

STARTTLS aims to protect the individual hops between SMTP servers, primarily protecting messages from passive eavesdroppers. As we will discuss in Section 3, STARTTLS is typically not used to authenticate destination mail servers, but rather provides opportunistic encryption (unlike the TLS deployment for HTTPS). In almost all cases, mail servers do not validate presented certificates and will relay messages over cleartext if STARTTLS is not supported. Relays still have access to messages and can freely read and modify message content.

The STARTTLS RFC does not define how clients should validate presented certificates. While the RFC suggests that the recipient's domain (e.g., `gmail.com`) be present in the certificate, it also permits checking the fully qualified domain name (FQDN) of the MX server. This removes the need for third-party mail servers (e.g., shared hosting like Google Apps for Work) to present a trusted certificate for each hosted domain. However, it also enables network-level attackers to falsely report MX records that point to an attacker-controlled domain. Without additional security add-ons (such as DANE [12]), this attack remains a real threat.

### 2.2 Authenticating Mail

Mail servers deploy several mechanisms for authenticating and verifying the integrity of received mail, including SPF, DKIM, and DMARC. While STARTTLS protects individual hops between SMTP servers, these protocols allow recipients to verify that messages do not have a spoofed sender and further provide a mechanism to report forged messages. A more detailed discussion of each protocol and shortcomings is available from MAAWG [10]. We describe the interplay between each system in Figure 2.

**DKIM** DomainKeys Identified Mail (DKIM) allows SMTP servers to detect if incoming messages have been spoofed or modified during transit (RFC 6376 [9]). In order to utilize DKIM, a sender appends the `DKIM-Signature` field to the message header (❷). This header contains a digital signature of the message that is associated with the domain name of the sender. Upon message

delivery, the recipient can retrieve the sender's public key through a DNS request, and verify the message's signature. DKIM does not specify what action the recipient should take if they receive a message with an invalid or missing cryptographic signature. Instead the organization must have a predetermined agreement with the sender.

**SPF** Sender Policy Framework (SPF) allows an organization to publish which hosts are authorized to send mail for their domain (RFC 7208 [23]). To deploy SPF, gthe organization publishes a DNS record that specifies which hosts or CIDR blocks belong to the organization. Upon receiving mail, the recipient performs a DNS query to check for an SPF policy and can choose to reject emails that do not originate from the specified servers. SPF further allows organizations to delegate a portion or the entirety of their SPF policy to another organization.

**DMARC** Domain-based Message Authentication, Reporting, and Conformance (DMARC) builds upon DKIM and SPF and allows senders to suggest a policy for authenticating received mail (RFC 7489 [25]). Senders publish a DNS TXT record (named `_dmarc.domain.com`) that indicates whether the sender supports email authentication (i.e., DKIM and/or SPF), and what action recipients should take if authentication fails (e.g., invalid DKIM signature) or if no signature is present. DMARC further allows organizations to request daily aggregate reports on spoofed messages that other servers receive.

# 3. CONFIDENTIALITY IN PRACTICE

To understand how mail confidentiality is protected in practice, we measured STARTTLS adoption from two perspectives: the protection of emails sent to/from Gmail and the configuration of SMTP servers associated with the Alexa Top 1 Million Domains.

## 3.1 Gmail

On weekdays between April 1 and April 26, 2015, Gmail was able to initiate STARTTLS connections for 79.8% of outgoing messages, and 53.7% of incoming connections initiated a STARTTLS session*. While the number of protected messages is consistent between weekdays during this period, there is an average 7.2% increase in the number of inbound connections that initiate a TLS session during U.S. weekends. During weekends, 57.6% of incoming connections and 79.9% of outgoing connections used STARTTLS. This may be because more personal and less business email is sent on weekends and personal accounts tend to be provided by large webmail providers (e.g., Gmail and Yahoo Mail).

We analyzed the cipher suites chosen by incoming Gmail connections on April 30, 2015 and found that 84.2% of TLS connections (45.2% of all incoming connections) chose a perfect forward secret cipher suite. 51.703% used AES-128-GCM, 45.643% used RC4, and 2.746% used AES-128 (Table 3).

STARTTLS adoption has been consistent, though growth is slow with a few exceptions. There has been a 54% increase (52% to 80%) in outbound and an 84% increase (32.5% to 59.9%) in inbound connections that utilize STARTTLS between January 2014 and April 2015. As can be seen in Figure 3, there are two areas of immediate interest. Between May 10 and May 30 2014, the outbound STARTTLS jumped from 47% to 71%. This was likely due to Yahoo and Microsoft (outlook.com) deploying STARTTLS. Second, between Oct 8 and Oct 17, outbound STARTTLS dropped from 73% to a low of 50%. The lowest point occurred on October 14, which corresponds with the public disclosure of the POODLE vulnerability [13], and the drop may be due to erroneous misconfigurations attempting to disable SSLv3.

*This excludes messages marked as spam. Data is available at http://www.google.com/transparencyreport/saferemail.

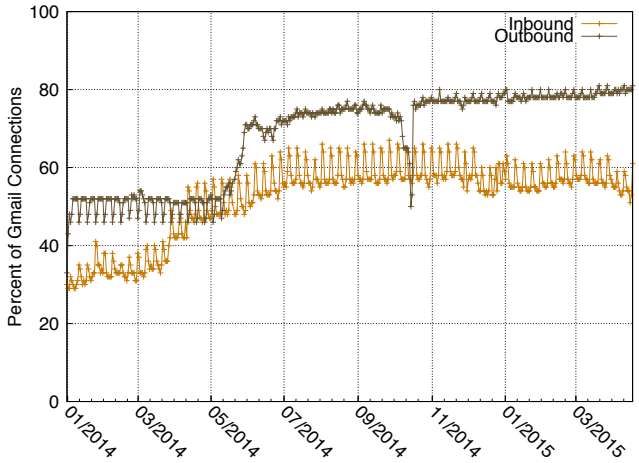


Figure 3: **Historical Gmail STARTTLS Support** — Inbound connections that utilize STARTTLS increased from 33% to 60% for weekdays between January 2014 and April 2015. Weekends consistently have close to 10% more connections that support STARTTLS than weekdays. Support for outgoing STARTTLS increased from 52% to 80% during this period.

In order to understand how our measurements are biased towards Gmail, we compared our measurements with previous estimates from Facebook. During May 2014, Facebook, which sends email notifications, found that 76% of destination mail servers supported STARTTLS and 58% of notifications were encrypted [17]. During this period, only 47% of outgoing Gmail messages were protected by STARTTLS. By August 2014, 95% of Facebook notifications were protected during transit by STARTTLS, after several large webmail providers, notably Microsoft and Yahoo, deployed STARTTLS [18]. During this time period, STARTTLS protected outbound Gmail connections also increased from 47% to 74%.

Despite the same opportunistic STARTTLS policy, Gmail has a generally lower percentage of outgoing mail protected by STARTTLS than Facebook, likely because Facebook's destination addresses are mostly personal accounts, which are biased towards large webmail providers (e.g., Gmail, Yahoo Mail, Outlook.com).

The mail transited by Gmail provides one perspective on mail security. However, a large percentage of messages are transited to/from large providers and the breakdown of messages does not necessarily represent how well other organizations have deployed STARTTLS. For example, of the 877 domains that Gmail transited mail to on April 26, 2015, only 58% of the domains accepted 100% of messages over TLS and of the 26,406 domains that transited mail to Gmail, only 29% protected 100% of messages with STARTTLS. In order to understand the complete picture, it is also important to consider how smaller organizations have deployed STARTTLS.

## 3.2 Organizational Deployment

To measure how organizations have deployed STARTTLS, we analyzed the configuration of mail servers associated with the Alexa Top 1 Million domains [2]. We queried the MX records for the Top 1 Million domains on April 26, 2015 from the University of Michigan and attempted an SMTP and STARTTLS handshake using ZMap [16]. We discuss the ethical implications of performing active scanning and provide details on how we reduce scan impact in our previous work [14, 16].

| Status | Top 1M Domains | |
|---|---|---|
| No MX records | 152,944 | (15.29%) |
| No resolvable MX hostnames | 11,967 | (1.20%) |
| No responding SMTP servers | 49,125 | (4.91%) |
| SMTP Server | 792,494 | (79.2%) |
| SMTP Server—No STARTTLS support | 144,464 | (18.2%) |
| SMTP Server—STARTTLS support | 648,030 | (81.8%) |

Table 1: **Top 1 Million Domains Scan Results** — We measured STARTTLS deployment for SMTP servers among Alexa Top 1M domains. 81.8% of SMTP servers support STARTTLS.

| | Matches Domain | Matches Server | Matches Neither |
|---|---|---|---|
| Trusted | 4,602 (0.6%) | 270,723 (34.2%) | 143,113 (18.1%) |
| Untrusted | 4,345 (0.6%) | 21,057 (2.7%) | 181,242 (22.9%) |
| **Total** | **8,947 (1.1%)** | **291,780 (36.8%)** | **324,355 (40.98%)** |

Table 2: **Certificates for Top 1 Million Domains** — While 52% of domains' SMTP servers present trusted certificates, only 34.2% of trusted certificates match the MX server, and only 0.6% are valid for the recipient domain.

| TLS Version | Key Exchange | Symmetric Cipher | HMAC | Inbound Traffic |
|---|---|---|---|---|
| TLSv1.2 | ECDHE | AES-128-GCM | SHA-256 | 51.500% |
| TLSv1 | ECDHE | RC4 | SHA-1 | 29.225% |
| TLSv1 | RSA | RC4 | SHA-1 | 14.403% |
| TLSv1.2 | ECDHE | AES-128 | SHA-1 | 1.586% |
| TLSv1.2 | RSA | RC4 | SHA-1 | 1.147% |
| TLSv1 | ECDHE | AES-128 | SHA-1 | 0.999% |
| TLSv1.1 | ECDHE | RC4 | SHA-1 | 0.723% |
| TLSv1.2 | RSA | AES-128-GCM | SHA-256 | 0.203% |
| SSLv3 | RSA | RC4 | SHA-1 | 0.060% |
| TLSv1.2 | ECDHE | RC4 | SHA-1 | 0.060% |
| TLSv1 | RSA | AES-128 | SHA-1 | 0.050% |
| TLSv1.1 | RSA | RC4 | SHA-1 | 0.024% |
| TLSv1.1 | ECDHE | AES-128 | SHA-1 | 0.011% |
| TLSv1.1 | ECDHE | AES-256 | SHA-1 | 0.004% |
| TLSv1.2 | RSA | AES-256 | SHA-1 | 0.003% |
| TLSv1.2 | RSA | AES-128 | SHA-1 | 0.001% |
| TLSv1 | RSA | RC4 | MD5 | 0.001% |

Table 3: **Cipher Suites for Inbound Gmail Traffic** — 80% of inbound Gmail connections are protected by TLS. Here, we present the selected cipher suites for April 30, 2015.

| Mail Provider | Domains | STARTTLS | Trusted Certificate | Certificate Matches |
|---|---|---|---|---|
| Gmail | 126,419 (15.9%) | Yes | Yes | server |
| GoDaddy | 36,229 (4.6%) | Yes | Yes | server |
| Yandex | 12,326 (1.6%) | Yes | Yes | server |
| QQ | 11,295 (1.4%) | Yes | Yes | server |
| OVH | 8,508 (1.1%) | Yes | Yes | mismatch |
| Other | 597,717 (75.4%) | – | – | – |

Table 4: **Top Mail Providers for Top 1 Million Domains** — Five mail providers are used by 25% of Top 1M domains for mail transport. All five support STARTTLS for incoming mail.

We were able to connect to valid SMTP servers for 792,494 domains[2] (Table 1). A large number of domains share the same mail servers—only 276,337 mail servers were specified by the 841,619 domains with MX records. Many of these domains are hosted by large mail providers and five providers service email for 24.6% of domains (Table 4).

648,030 (81.8%) of mail-enabled domains supported STARTTLS. Only 5 domains within the Top 50 did not: `wikipedia.org`, `vk.com`, `weibo.com`, `yahoo.co.jp`, and `360.cn`. With the exception of two sites, all presented certificates with RSA keys; 10.0% used 1024-bit keys, 86.4% of domains used 2048-bit keys, and 3% used 4096-bit or larger keys. Only 316 domains presented 512-bit RSA certificates. 25.3% of domains supported perfect toward secrecy end completed an ephemeral Diffie-Hellman key exchange. 59.2% of domaisn used RC4 and 40.8% used AES; 25 sites selected 3DES. In summary, most sites that deployed STARTTLS deployed secure certificates. However, similar to the HTTPS ecosystem, sites are slow in deploying modern, secure cipher suites.

### *3.2.1 STARTTLS Certificates*

As part of the STARTTLS handshake, each mail server presents an X.509 certificate. While RFC 3207 [20] suggests that certificates match the mail domain (e.g., `gmail.com`), it also permits certificates that only match the MX server itself (e.g. `aspmx.l.google.com`). However, certificates that match the MX server do not provide true authentication unless the MX records for the domain are cryptographically signed. Otherwise, an active attacker can return the names of alternate, attacker controlled, MX servers in the initial MX query. Realistically operators cannot rely on this today—recent studies have found that less than 0.6% of `.com` and `.net` domains have deployed DNSSEC [34].

In our scan, 414,374 domains (52% of domains with valid SMTP servers and 64% of domains that supported STARTTLS) presented certificates that validated against the Mozilla NSS root store [28] (Table 2). However, only 0.6% of domains presented certificates that matched the domain and 34.2% of sites present trusted certificates that match the MX server. Surprisingly, 18.1% of domains present trusted certificates that match neither.

The large number of CA-signed, but mismatched, certificates are primarily due to several mail hosting providers, including `psmtp.com` and `pphosted.com`, who are incorrectly using wildcald certificates. In the remaining cases, certificates were simply for different domains. 33,281 domains presented expired cretificates, 60 domains were signed by untrusted CAs and 55 certificates were invalidly signed by a parent certificate whose type mismatched the child certificate.

## 3.3 Common Software Implementations

In order to understand why such a large number organizations have not deployed STARTTLS and why only half of inbound connections to Gmail initiate a STARTTLS connection, we investigated the five five most popular SMTP implementations, which account for 97% of identifiable mail servers for the top million domains (Table 5). We specifically tested whether each implementation initiated STARTTLS connections, supported STARTTLS for incoming connections, and how the implementation validated certificates. We installed the latest version of each SMTP server on an Ubuntu

[2]We note that while the 15.3% of domains missing MX records initially appears high, secondary scans confirm these results, and the domains missing MX records (e.g., `t.co`, `googleusercontent.com`, `blogspot.com`) are consistent with domains that do not need mail servers.

14.04.1 LTS system, except for Microsoft Exchange, which was readily documented online [27].

By default, Microsoft Exchange, Exim, and Sendmail initiate STARTTLS connections when delivering messages; Postfix and qmail—which together account for nearly 35% of all identifiable mail servers on the public IPv4 address space—send all messages over cleartext unless explicitly configured to use STARTTLS. All of the servers we tested fail to open and send mail in cleartext if STARTTLS is not available. Postfix and Microsoft Exchange Server support inbound STARTTLS connections without user intervention by generating a self-signed certificate on install. The remaining servers do not accept TLS connections without manual configuration. Postfix and Exchange, which provide confidentiality by default, account for 22% of the servers for the top million domains.

Postfix was the only server capable of performing both server-based and domain-based certificate validation, although its documentation specificially recommends *against* enabling validation when interacting with the greater Internet [33]. Exim, qmail, Sendmail, and Microsoft Exchange do not support validating the destination domain when relaying mail.

## 3.4 Popular Mail Providers

Given that a large percentage of mail is transited through a small number of popular providers, a single change can have a large impact on entire ecosystem, as is demonstrated in Figure 3. We measured inbound and outbound STARTTLS support for 19 common webmail providers and Internet service providers (Table 6). Only one provider—Lycos—did not support inbound STARTTLS. Two providers—facebookmail and OVH—presented certificates that matched neither their domain or the hostname of their MX server. *None* of the providers presented a certificate that matched the their domain and would have allowed authentcation. Less than half of the providers negotiated perfect forward secret cipher suite.

When sending mail, three of the providers—Lycos, GoDaddy, and OVH—did not initiate STARTTLS connections; the remaining providers initiated STARTTLS connections, but did not validate certificates, effectively provided opportunistic encryption, but no authentication.

To test outgoing mail, we created an account on each provider and then sent mail to a Postfix server, which was configured to support STARTTLS using the self-signed certificate generated at install. To test incoming STARTTLS support, we connected to the mail servers listed in each domains' MX record and initiated a STARTTLS handshake.

## 3.5 Takeaways

There has been significant growth in STARTTLS adoption over the past year. However, much of this growth can be attributed to a handful of large providers, and smaller organizations continue to lag in deploying STARTTLS. As of March 2015, nearly half of inbound weekday connections to Gmail still fail to encrypt messages. This may partially be due to several popular SMTP implementations not starting STARTTLS connections by default.

All encryption is performed opportunistically—none of the providers nor implementations use TLS for authentication and only one common implementation supports validating a certificate against the destination domain. This may be in part due to the fact that the majority of certificates match neither the mail server nor the destination domain, and could be realistically used. Unfortunately until there is widespread STARTTLS support, organizations cannot realistically require encryption and until organizations deploy valid certificates, relays will be unable to perform stringent validation.

# 4. THREATS TO CONFIDENTIALITY

As deployed in practice, STARTTLS protects connections against passive eavesdroppers, but does not protect against active man-in-the-middle and man-on-the side attacks[3]. We examine two types of network attacks this enables. In the first, attackers take advantage of the fail open design of STARTTLS where SMTP severs fall back to cleartext if any errors occur during the STARTTLS handshake. This failure mode accommodates gradual industry adoption (as opposed to HTTPS which terminates the connection), but exposes STARTTLS to downgrade attacks, which can be easily accomplished by corrupting any packet during the STARTTLS handshake. In the second attack, falsified MX records via intercepted DNS requests or impersonated DNS authoritative zones enable a malicious SMTP server to falsely advertise themselves as the correct destination for messages. We estimate the prevalence of both attacks, keying in on the networks that tamper with STARTTLS negotiations and DNS servers reporting falsified MX records.

## 4.1 STARTTLS Prevention

An active attacker—or a legitimate organization with a vested interest in snooping email—can prevent email encryption by tampering with the establishment of a TLS session and causing it to fail open. In practice we observe two main ways this is performed: either the attacker mangles the packet containing the STARTTLS command, which prevents the TLS session from being initiated, or the attacker alters the server's EHLO response to remove STARTTLS from the list of server capabilities. If the client still persists on using STARTTLS, the attacker can then modify the packet containing the STARTTLS command to contain an invalid command. This causes the server to respond with an error stating that it does not support the (invalid) command.

An man-on-the-side attacker could also inject a packet containing an invalid command with the hopes that the injected packet will reach the server before the legitimate packet containing the STARTTLS command. The client's STARTTLS packet would then be dropped by the server's TCP stack as a duplicate packet, and the server would respond to the client rejecting the invalid command. In either case, the client interprets the error response as the server not supporting STARTTLS, and the connection will fail open to cleartext.

In order to measure the prevalence of STARTTLS being stripped from connections, we performed a TCP SYN scan of the public IPv4 address space on port 25 and attempted to perform an SMTP and STARTTLS handshake with responsive hosts, regardless of the capabilities advertised in the `EHLO` response. The scan was performed on April 20, 2015 from the University of Michigan campus. 14% of SMTP servers echo back the received command when replying with an *Invalid Command* error. Using these responses, we can infer whether the STARTTLS command was altered in transit, and investigate if and how our commands are being replaced.

Our scan found 14.1M hosts with port 25 open, 8.9M SMTP servers, and 4.6M SMTP servers that support STARTTLS (Table 9). Of the 4.2M hosts that did not complete a TLS handshake, 623,635 echoed back the received command. We classified the echoed back responses and found that 617,093 (98.95%) indicated `STARTTLS` (and indeed did not support STARTTLS), 5,750 (0.92%) returned `XXXXXXXX`, 786 (0.14%) responded with `STAR` or `TTLS`, and 6 responded with `BLUF`.

The `STAR` and `TTLS` commands are four character command truncations and are likely not due to an attack. Prior to ESTMP, SMTP

[3] A *man-on-the-side attack* is an active attack in which an attacker can read traffic and inject messages, but cannot modify or delete messages sent between the two parties.

| Mail Software | Top 1M Domains Market Share | Public IPv4 Market Share | STARTTLS Incoming | StartTLS Outgoing | Server Validation | Domain Validation | Reject Invalid Certificates | TLS Version |
|---|---|---|---|---|---|---|---|---|
| exim 4.82 | 34% | 24% | ◐ | ● | ○ | ○ | ○ | 1.2 |
| Postfix 2.11.0 | 18% | 21% | ● | ◐ | ◐ | ◐ | ◐ | 1.2 |
| qmail 1.06 | 6% | 1% | ◐ | ◐ | ○ | ○ | ○ | 1.2 |
| sendmail 8.14.4 | 5% | 4% | ◐ | ● | ○ | ○ | ○ | 1.2 |
| Exchange 2013 | 4% | 12% | ● | ● | ◐ | ○ | ◐ | 1.0 |
| Other | 3% | <1% | | | | | | |
| Unknown | 30% | 38% | | ● default behavior ◐ non-default ○ no support | | | | |

Table 5: **Popular Mail Transfer Agents** — We investigated the default behavior for five popular MTAs. By default, Postfix and qmail do not initiate STARTTLS connections. All five MTAs we tested will fall back to cleartext if the STARTTLS connection fails.

| Provider | Incoming TLS Version | Incoming Key Exchange | Incoming Cipher | Certificate Matches | Outgoing TLS Version | Outgoing Key Exchange | Outgoing Cipher |
|---|---|---|---|---|---|---|---|
| Gmail | 1.2 | ECDHE | AES-128-GCM | server | 1.2 | ECDHE | AES-128-GCM |
| Yahoo | 1.2 | ECDHE | AES-128-GCM | server | 1.0 | ECDHE | RC4-128 |
| Outlook | 1.2 | ECDHE | AES-256-CBC | server | 1.2 | ECHDE | AES-256 |
| iCloud | 1.2 | ECDHE | AES-128-GCM | server | 1.2 | DHE | AES-128-GCM |
| Hushmail | 1.2 | RSA | RC4-128 | server | 1.2 | ECDHE | AES-256-GCM |
| Lycos | – | – | – | – | – | – | – |
| Mail.com | 1.2 | ECHDE | AES-256-CBC | server | 1.2 | DHE | AES-256-GCM |
| Zoho | 1.0 | RSA | RC4-128 | server | 1.0 | RSA | RC4-128 |
| Mail.ru | 1.2 | RSA | RC4-128 | server | 1.2 | ECDHE | AES-256-GCM |
| AOL | 1.0 | RSA | RC4-128 | server | 1.0 | DHE | AES-256-CBC |
| QQ | 1.1 | RSA | RC4-128 | server | 1.0 | DHE | AES-256-CBC |
| Me.com | 1.2 | ECHDE | AES-128-GCM | server | 1.2 | DHE | AES-128-GCM |
| facebookmail | 1.0 | RSA | AES-128-CBC | mismatch | 1.0 | ECDHE | AES-128 |
| GoDaddy | 1.2 | RSA | RC4-128 | server | – | – | – |
| Yandex | 1.2 | RSA | AES-128-GCM | server | 1.2 | ECDHE | AES-256-CBC |
| OVH | 1.2 | RSA | AES-128-GCM | mismatch | – | – | – |
| Comcast | 1.2 | RSA | RC4-128 | server | 1.2 | DHE | AES-128-CBC |
| AT&T | 1.2 | ECDHE | AES-128-GCM | server | 1.0 | ECDHE | RC4-128 |
| Verizon | 1.2 | RSA | AES-128-GCM | server | 1.0 | DHE | AES-128-CBC |

Table 6: **Encryption Behavior of Mail Providers** — We measured support for incoming and outgoing STARTTLS among various popular mail providers. While most providers supported STARTTLS, *none of them* validated our certificate, which was self-signed.

| Provider | Servers Providing Invalid MX Answers | Servers Providing Invalid IP Answers | Unique Invalid MX Servers | Unique Invalid IPs | Responsive Invalid Mail Servers |
|---|---|---|---|---|---|
| Gmail | 30,931 | 23,134 | 146 | 1,150 | 144 |
| Yahoo | 31,219 | 55,459 | 130 | 1,117 | 114 |
| Outlook.com | 29,618 | 23,145 | 117 | 1,059 | 110 |
| Mail.ru | 31,214 | 25,796 | 97 | 1,053 | 110 |
| QQ | 30,091 | 55,467 | 122 | 1,171 | 111 |

Table 7: **Falsified DNS Responses** — We scanned the public IPv4 address space for DNS servers that returned falsified MX records and IP addresses for five popular mail providers. This data excludes loopback addresses and obvious configuration errors.

| | Nov. 2013 | April 2015 | Change |
|---|---|---|---|
| **Overall failure rate** | **10.65%** | **6.14%** | **−4.42%** |
| Crypto failures: | | | |
| Weak crypto key (<1024 bits) | 21.00% | 15.08% | −5.92% |
| Key is revoked | 0.02% | 0.01% | −0.01% |
| Signature algorithm not supported | 0.27% | 0.26% | −0.02% |
| Key is expired | | 0.06% | |
| Body hash doesn't match signature | | 18.66% | |
| Protocol version incorrect | 0.59% | 3.32% | +2.73% |
| Some DKIM tags are duplicated | | 0.05% | |
| Other error | 77.91% | 62.55% | −15.36% |

Table 8: **Gmail DKIM Errors** — We present the breakdown of Gmail DKIM validation failures for Nov. 2013 and April 2015.

| Scan Result | Hosts |
|---|---|
| TCP port 25 open | 14,131,936 |
| Responsive SMTP server | 8,850,664 |
| Successful STARTTLS handshake | 4,620,561 |

Table 9: **IPv4 SMTP Scan Results** — We could handshake STARTTLS with 52% of the SMTP servers our scans identified.

| | Hosts | |
|---|---|---|
| Command not echoed | 3,606,468 | (85.26%) |
| `STARTTLS` echoed correctly | 617,093 | (14.59%) |
| `STARTTLS` replaced | 5,756 | (0.14%) |
| Command truncated to four characters | 786 | (0.02%) |

Table 10: **STARTTLS Manipulation** — We could extract an echoed command from 14.75% of servers that sent errors in response to `STARTTLS`. 0.14% of these responses indicate that the command was manipulated before reaching the server.

| | Top 1M Domains | IPv4 Hosts |
|---|---|---|
| Cisco tampering | 2,563 | 41,405 |
| BLUF tampering | 0 | 6 |

Table 11: **Prevalence of STARTTLS Stripping** — We find evidence of STARTTLS being stripped from SMTP connections by Cisco security devices.

| Type | ASes | |
|---|---|---|
| Corporation | 182 | (43.0%) |
| ISP | 74 | (17.5%) |
| Financial | 57 | (13.5%) |
| Academic | 35 | (8.3%) |
| Government | 30 | (7.1%) |
| Healthcare | 14 | (3.3%) |
| Unknown | 12 | (2.8%) |
| Airport | 9 | (2.1%) |
| Hosting | 7 | (1.7%) |
| NGO | 3 | (0.7%) |

Table 12: **ASes Stripping STARTTLS** — We categorize the ASes for which 100% of SMTP servers showed behavior consistent with STARTTLS stripping.

| | Hosts |
|---|---|
| DNS servers | 13,766,099 |
| Responsive DNS servers | 8,860,639 |
| Any invalid MX responses | 234,756 |
| Class of invalid behavior: | |
| Identical response regardless of request | 131,898 |
| Returns loopback address | 16,015 |
| Returns private network address | 7,680 |
| Flipped bits in response | 56,317 |
| Falsified DNS record | 178,439 |

Table 13: **Invalid or Falsified MX Records** — We scanned the IPv4 address space for DNS servers that provided incorrect entries for the MX servers for five popular mail providers.

| | | | |
|---|---|---|---|
| Tunisia | 96.13% | Reunion | 9.28% |
| Iraq | 25.61% | Belize | 7.65% |
| Papua New Guinea | 25.00% | Uzbekistan | 6.93% |
| Nepal | 24.29% | Bosnia and Herzegovina | 6.50% |
| Kenya | 24.13% | Togo | 5.45% |
| Uganda | 23.28% | Barbados | 5.28% |
| Lesotho | 20.25% | Swaziland | 4.62% |
| Sierra Leone | 13.41% | Denmark | 3.69% |
| New Caledonia | 10.13% | Nigeria | 3.64% |
| Zambia | 9.98% | Serbia | 3.11% |

Table 14: **Countries Affected By STARTTLS Stripping** — We measure the fraction of incoming of Gmail messages that originate from the IPs we find stripping TLS from SMTP connections. Here, we show countries with the most mail affected by STARTTLS stripping and the affected percentage of each country's incoming mail between April 20–27, 2015.

| | |
|---|---|
| Slovakia | 0.08% |
| Romania | 0.04% |
| Bulgaria | 0.03% |
| India | 0.02% |
| Israel | 0.01% |
| Switzerland | 0.01% |
| Poland | 0.01% |
| Ukraine | 0.01% |

Table 15: **Countries Affected By Falsified DNS Records** — We measure the fraction of mail received by Gmail on May 21, 2015 from the IP addresses pointed to by false Gmail DNS entries. Here, we show the breakdown of mail from each country that originates from one of these addresses for the countries with the most affected mail.

| Provider | SPF Policy | DMARC Policy |
|---|---|---|
| Gmail | soft fail | none |
| Yahoo | neutral | reject |
| Outlook | soft fail | none |
| iCloud | soft fail | none |
| Hushmail | soft fail | – |
| Lycos | soft fail | – |
| Mail.com | fail | – |
| Zoho | soft fail | – |
| Mail.ru | soft fail | none |
| AOL | soft fail | reject |
| QQ | soft fail | none |
| Me.com | soft fail | none |
| Facebook | fail | reject |
| GoDaddy | fail | none |
| Yandex | soft fail | – |
| OVH | neutral | – |
| Comcast | neutral | none |
| AT&T | – | – |
| Verizon | neutral | – |

Table 16: **SPF and DMARC Policies** — The majority of popular mail providers we tested posted an SPF record, but only three used the "strict fail" policy. Even fewer providers posted a DMARC policy, of which only three used "strict reject."

commands were all four characters and we were able to confirm that any sent command was truncated to four characters in older software. However, the `XXXXXXXX` and `BLUF` commands were due to the STARTTLS command being altered. We summarize the prevalence of those two types of stripping behavior in Table 10.

The `XXXXXXXX` replacements are potentially attributed to security products intercepting and stripping the STARTTLS command. In one prominant example, Cisco Adaptive Security Appliances (ASA) [5] and Cisco IOS Firewall [6] which both support replacing the STARTTLS command with Xs in order to facilitate the mail inspection as part of their *inspect smtp* and *inspect esmtp* configurations.

Cisco SMTP inspection requires access to plaintext traffic in order to restrict commands, stop potentially malicious traffic, and audit transmitted mail. Cisco advertises that their products are specifically capable of searching for and dropping messages with invalid characters in email addresses, invalid SMTP commands, and long commands that may be attempting to exploit buffer overflows [7].

We are unable to attribute the `BLUF` replacement to any commonly known security software. The six hosts affected by this replacement also had the `PIPELINING` and `CHUNKING` capabilities in the `EHLO` response masked to `HIPELINING` and `PHUNKING`, respectively. No hosts besides these six showed this type of behavior, and all six were located in Ukraine.

We found 5,756 servers that have the STARTTLS command corrupted. However, this is an understimate of the total affected servers, because we can only detect this behavior when servers echo back the received command—approximately 14% of SMTP servers. We were able to increase our dataset coverage by leveraging the fact that Cisco SMTP inspection boxes and potentially other vendors also replace the STARTTLS text that is sent as part of the `EHLO` response sent by server that support STARTTLS. Looking for those X'ed out STARTTLS capabilities in `EHLO` response allowed us to find 35,649 additional servers that perform TLS downgrading. Overall, we did find 41,405 servers that support STARTTLS, but cannot complete a handshake due to handshake corruption.

### *4.1.1 Prevalence*

The 41,405 STARTTLS-stripped SMTP servers belong to 4,714 ASes (15% of all ASes with an SMTP server) and are located in 191 countries (86% of countries with SMTP servers). These servers provide mail for 2,563 domains in the Top 1 Million. In 423 ASes (736 hosts), 100% of SMTP servers were affected by STARTTLS stripping. The AS performing stripping on 100% of the inbound and outbound email with the most SMTP servers (21) belonged to Starwood Hotels and Resorts (AS 13401). The classification of the ASes with 100% stripping is shown in Table 12. Overall, no single demographic stands out—the distribution is spread over networks owned by governments, Internet service providers, corporations, and financial, academic, and health care institutions. We note that several airports and airlines appear on the list, including an AS belonging to a subsidiary of Boingo (AS 10245), a common provider of in-flight and airport WiFi.

To understand the amount of email traffic affected by TLS blocking, we measured the amount of email traffic transited to/from these devices from Gmail's perspective. While the overall percentage of affected mail is low, a hanful of countries have a high stripping rate. For example, 96.13% of mail transited from Tunisia to Gmail is affected by STARTTLS stripping. 9 countries experience over 10% stripping, and 16 experience more than 5% stripping (Table 14).

It is important to note that the devices that are stripping TLS from SMTP connections are not inherently malicious, and many of these devices may been deployed to facilitate legitimate filtering. However, regardless, methodology results in messages being sent in cleartext over the public Internet, enabling passive eavesdropping and other attacks. Furthermore, the Cisco documentation does not outline the downsides of this methodology, and administrators may not be aware that the setting puts users at risk. Instead of stripping TLS, manufacturers should consider deploying in-line devices that accept and initiate STARTTLS connections, but inspect messages, before forwarding them to an internal mail server.

## 4.2 DNS Hijacking

***[TODO: redo with new numbers and experiment]***

A second method for intercepting mail is to spoof the DNS records for the destination SMTP server. We investigated whether DNS servers are providing false MX records for gmail.com or false IP addresses for any of gmail's MX servers. IP addresses or MX records for gmail.com, yahoo.com, outlook.com, qq.com (popular Chinese webmail provider), and mail.ru (popular Russian webmail provider) and find that 178,439 of 8,860,639 (2.01%) of publicly accessible DNS servers provided invalid IPs or MX records (Table 13).

We searched for spoofed servers by implementing a DNS scanner for ZMap and performing ten ZMap [16] scans of the publicly accessible IPv4 address space on April 25, 2015. For each of the five domains (gmail.com, yahoo.com, outlook.com, qq.com, and mail.ru), we queried both the MX record for the domain and the A record for the highest priority MX server. 13.8 million servers responded with a valid DNS responses; 8.9 million servers resolved one or more of the queries(Table 13). From these ten scans, we generated a single set of all IP addresses that responded with invalid addresses or MX servers (235K hosts).

We performed follow-up DNS queries against these 235K hosts, re-querying the MX record along with an A query for each of responses from the MX query, the valid MX servers for each domain, umich.edu, and doesnotexist.umich.edu (a non-existent domain). 56K hosts provided correct results during this secondary scan and appear to have produced invalid responses due to bit-flips or other packet corruption. 132K hosts responded to all queries with the same publicly accessible address (including the non-existent domain), 7,680 hosts responded to all queries with a reserved or private address (e.g., 10.0.0.0/8), and 16,015 hosts responded with a loopback address (e.g., 127.0.0.1).

### *4.2.1 Population*

We further investigated the 31,774 DNS servers that provided incorrect answers for gmail.com and associated MX servers, but did not blindly return the same results for all DNS queries. 17,216 (54%) of the DNS hosts do not provide any incorrect records, but were instead missing one or more of the MX servers. The remaining 14,558 hosts returned 1,150 unique invalid IP addresses (IPs that were not in a Google controlled AS) and were located in 521 ASes. Of those 1150 only 144 (12.5%) of these hosts completed an initial SMTP handshake.

83.6% of these hosts were located in five ASes: 62% from Unified Layer (American Hosting Provider), 11.7% from ChinaNet, 5.3% Telecom Italia (Italian ISP), 2.4% from SoftLayer Technologies, 2.0% from eNom. In the case of Unified Layer, 9,073 hosts all pointed back to seven unique servers within the AS; two of the servers accepted public SMTP connections. Both ran Exim 4.82. In the case of the ChinaNet, hosts pointed to a local address, private subnet, or one of 42 servers in the ChinaNet As; one completed an SMTP handshake.

The devices in the Telecom Italia AS all returned seemingly random IP addresses within the subnet 198.18.1.0/24. Further investigation found that these devices returned monotonically increasing

IPs within that range for all DNS queries; because the resulting IPs were different for each query they were not caught in earlier checks; none of the destination IPs appeared to be SMTP servers. The SoftLayer hosts all responded with one of eighteen servers; ten completed SMTP handshakes. All eNom hosts pointed to a single IP, which did not accept SMTP connections. The remaining 2,386 servers were located in 533 ASes and 69 countries. 144 of the unique destination IPs successfully completed an SMTP handshake.

Our results show that there are a large number of ASes where DNS servers provide false records for Gmail's MX servers. However, it is difficult to tell whether these falsified records are malicious, or are due to other misconfiguration. The number of hosts affected is not immediately clear by falsified DNS records is not immediately clear. Standard practice dictates that DNS server should not perform recursive resolution (therefore, not providing our client with the MX records for Gmail.com). Second, it is unclear how many hosts may resolve queries against the given DNS server.

In order to understand whether mail is being intercepted by these hosts, we measured the amount of mail Gmail received from the addresses listed in the falsified DNS records. As shown in Table 15, only a small percentage of mail originates from these hosts and a large majority are sending spam.

# 5. AUTHENTICATION IN PRACTICE

While STARTTLS protect messages against passive eavesdropping, it does not provide authentication—mail can be modified or spoofed altogether. As described in Section 2, SPF, DKIM, and DMARC have been developed to authenticate incoming mail. In this section, we describe how these protocols have been deployed in practice.

## 5.1 SPF

SPF enables recipients to detect whether messages were sent by an authorized server by checking for the sender's IP address in the organization's SPF policy—a special DNS TXT record.

Only 401,356 domains—47% of the top 1M domains with MX records—have published SPF policies. Of theses, 86,919 (21.7%) have hard fail policies (email outside of the specified networks should be rejected), 232,736 (58.0%) of domains have soft fail policies (email should be accepted, but marked as suspect), and 81,701 (20.3%) of domains have no set policy (Table 19).

While SPF policies should only include the IP addresses of the organization's SMTP servers, more than 133,490 (60.9%) allow CIDR ranges larger than /24. 99,698 (29.2%) specify CIDR ranges larger than or equal to a /16, and 1,333 (0.4%) specify more than a /8 worth of addresses. 1,293 domains specified more than $2^{24}$ IP addresses (the size of a /8) in their SPF policies. Of these, only 104 of these did not point to 10.0.0.0/8 and 62 domains appeared to be blatant misconfigurations (e.g., 255.255.255.255/8). Of the 32 remaining domains, 13 were Apple-owned companies that included Apple's IANA-assigned /8 (17.0.0.0/8). The remaining 20 domains incorrectly allowed larger subnets to send mail.

Many domains point their SPF records at cloud hosting providers. 3.24% of domains with SPF records contain an entry that points at IP-space used by Amazon EC2. EC2 assigns addresses randomly on boot from a pool of public addresses—any customer may be adjacent to any other customer. Similarly, 16.0% have entries pointing towards Softlayer (IBM Cloud Services), and 8.1% point towards Azure (Microsoft Cloud).

An SPF policy can also restrict emails to a list of FQDNs or to hosts with reverse DNS entries that point to the sender's domains. To understand how SPF in practice, we investigated the SPF records deployed by the top 1 million domains. 255,867 domains specify to allow mail from their MX servers, and 104 domains allow mail from servers with reverse DNS names that match the domain. 10,432 domains redirect (or fully delegate their SPF policy) to another provider and 213,464 (53.2%) domains include records from one or more other domain's SPF policies.

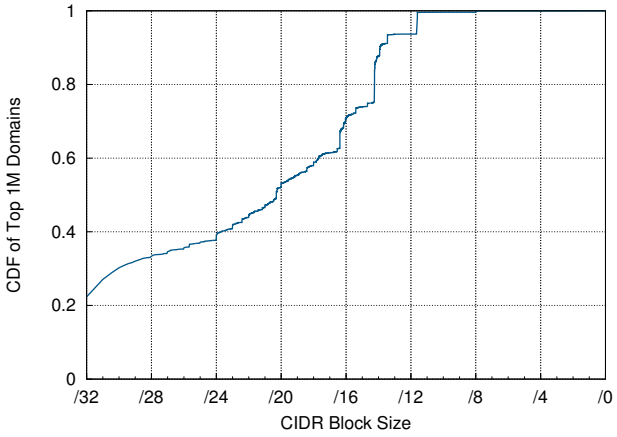


Figure 4: **Size of SPF Permitted Networks** — We show the CDF of number of addresses whitelisted in a full recursive resolution of the SPF records for the Top 1M domains.

While there could be potential for abuse if multiple organizations specified the same IP blocks, this is not the case for the most commonly included records. Instead, we found that domains commonly used these methods to include information from other mail providers. 68% redirected domains redirected to one of five domains: 35.7% redirected to Yandex (a Russian mail provider), 16.7% mailhostbox.com, 8.0% nicmail.ru, 3.9% serveriai.lt, and 3.5% to mail.ru. 3,813 (36.6%) of all redirects pointed to a Russian mail service. 136,473 domains (64% of domains with includes) included one of five large mail providers, Gmail (59,660), Outlook.com (44,216), websitewelcome.com (20,291), mandrillapp.com (16,606), and SendGrid (10,700).

Most mail providers publish SPF records with a soft fail policy—the exceptions being Facebook Mail, Mail.com, and GoDaddy, which all had hard fail policies. AT&T was the only provider we checked that did not have a valid SPF policy (Table 16).

From Gmail's perspective, 94% of messages received during April 2015 were authenticated using SPF; 0.42% could not be validated due to errors fetching the domain's SPF record (Table 21). However, despite the high percentage of validated mail, organizations are lagging in deploying SPF, with less than half of the domains we investigating publishing an SPF record. Further, organizations appear to be deploying SPF lazily—61% of the domains specified IP ranges larger than a /24 rather than specific e-mail servers and others specified IP ranges owned by shared cloud providers.

## 5.2 DKIM

DKIM allows a recipient to confirm that message content has not been altered during transit by validating an HMAC attached to the message. 83.0% of the messages received by Gmail in April 2015 had a DKIM signature, of which 6.14% failed validation. For 18.66% of the failed messages, validation failed due to the message signature not matching the message content (Table 8).

The DKIM failure rate itself fell by 4.42% between November 2013 and April 2015. The cryographic security of DKIM signatures has increased over time. In November 2013, 21% of DKIM failures

were caused by the use of a less-than 1024-bit signing key. By April 2015, only 15% of DKIM errors were caused by an insufficiently strong signing key.

Combined with SPF, Gmail was able to validate 94.40% of received messages in April 2015. 82.99% of messages had valid DKIM signatures; 92.42% of messages were sent from a permitted host or network by a domain with a published SPF ruleset. We show the breakdown of incoming mail authentication for Gmail in Table 17.

## 5.3 DMARC

DMARC allows organizations to post what action a recipient should take if a message fails DKIM or SPF validation. DMARC was recently introduced in 2015, but solves the real-world problem that organizations had no way to whether a message should have a DKIM signature, or what to do with messages that validation.

We measured organizational DMARC deployment by querying the DNS DMARC TXT record for each of the top 1 Million domains. Only 1.0% of domains publish DMARC policies; only 0.3% deploy policies to quarantine or reject messages missing a DKIM signature (Table 18). We list the DMARC policies for popular mail providers in Table 6.

While 98.9% of the Alexa Top 1 Million does not publish a DMARC policy, only 73.9% of incoming mail to Gmail comes from a domain without a DMARC policy, suggesting the security policies of large mail providers can have a large impact on overall mail security.

For domains that post DMARC policies, the most common policy is the empty policy, which allows mail that fails verification to continue to be transited, although the occurance is recorded and sent to a reporting address listed in the DMARC policy. Despite the most common policy of empty for Alexa domains, the most common policy on incoming mail to Gmail is reject, which accounts for 13.1% of all incoming mail, whereas the empty policy accounts for 11.7% of incoming mail. The least common policy in both cases was the quarantine policy, which suggests mail that fails verification be tagged or marked as spam.

Similarly, among the popular mail providers who post DMARC policies, the most common policy was the empty policy. Yahoo, AOL, and Facebook post DMARC reject policies. No provider used the quarantine policy (Table 16).

While DKIM has widespread deployment, with over 80% of incoming mail to Gmail signed using DKIM, only 26.1% of incoming mail has a DMARC policy. The lack of deployment of DMARC relative to DKIM limits the effectiveness of DKIM itself.

# 6. DISCUSSION

The mail community has retroactively applied several security measures to SMTP. 60% of incoming connections to Gmail are encrypted, and 94% of messages are authenticated with DKIM or SPF. In many ways, this is a feat, given that SMTP did not originally provide any support for transport security. However, in many ways, our two perspectives paint drastically different pictures of how mail security has been deployed. As can be seen by the 51% jump in encrypted inbound messages when Microsoft and Yahoo deployed STARTTLS, a large percentage of this success can be attributed to large mail providers who are pushing security forward. Unfortunately, as our scans for STARTTLS demonstrate, smaller organizations lag in deploying these mechanisms correctly—only 82% of domains have deployed STARTTLS and 65% of STARTTLS-enabled domains are not configured to allow senders to authenticate that they are connected to the correct mail server.

| Authentication Method | Nov. 2013 | April 2015 | Change |
|---|---|---|---|
| DKIM & SPF | 74.66% | 81.01% | +6.31% |
| DKIM only | 2.25% | 1.98% | −0.27% |
| SPF only | 14.44% | 11.41% | −2.99% |
| No authentication | 8.65% | 5.60% | −3.00% |

Table 17: **Gmail Incoming Mail Authentication** — During April 2015, 94.40% of incoming Gmail messages were authenticated with DKIM, SPF, or both.

| Published Policy | Gmail Messages | Top 1M Domains |
|---|---|---|
| Quarantine | 1.34% | 709 (0.09%) |
| Empty | 11.66% | 6,461 (0.82%) |
| Reject | 13.08% | 1,720 (0.22%) |
| Not published | 73.92% | 783,851 (98.9%) |

Table 18: **DMARC Policies** — We categorize DMARC policies for incoming Gmail messages from April 2015 and for Alexa Top 1M domains with MX records on April 26, 2015.

| Policy | Top 1M Domains | Recursive Top 1M |
|---|---|---|
| SPF Policy | 401,356 | 401,356 |
| Hard fail | 84,801 (21.13%) | 86,919 (21.65%) |
| Soft fail | 226,117 (56.34%) | 232,736 (57.99%) |
| Neutral | 80,394 (20.03%) | 81,701 (20.36%) |
| Redirect | 10,045 (2.50%) | 0 (0.00%) |

Table 19: **SPF Policies for Top 1M Domains** — We queried the SPF policies for the Alexa Top 1 Million domains for both the top-level record, and for full recursive resolution.

| Record Type | Top 1M Domains | Recursive Top 1M |
|---|---|---|
| IPv4 | 200,976 (33.08%) | 344,844 (40.22%) |
| IPv6 | 6,862 (1.13%) | 108,086 (12.61%) |
| A | 139,979 (23.04%) | 148,688 (17.34%) |
| MX | 249,345 (41.04%) | 255,867 (29.84%) |
| REDIRECT | 10,432 (1.72%) | 0 (0.00%) |

Table 20: **SPF Record Types for Top 1M Domains** — We show how hosts are whitelisted within an SPF record for both the top-level SPF record, and for full recursive resolution.

| SPF Policy | Gmail Messages |
|---|---|
| DNS timeout | <0.001% |
| Temporary error | 0.1840% |
| Permanent error | 0.1405% |
| Invalid record | 0.0978% |

Table 21: **SPF Errors for Incoming Gmail Traffic** — We show the breakdown of errors fetching SPF records for incoming mail. Temporary errors can be fixed by retrying later. Permenant errors mean the record was unable to be fetched.

While the state of mail delivery security is rapidly improving, there are several structural challenges the mail community needs to address in order to guarantee the confidentiality and integrity of mail delivery. In this section, we explore those challenges.

## 6.1 Challenges for Confidentiality

There are several major challenges for guaranteeing the confidentiality of mail in transit. First, unlike HTTPS, which has HSTS, there is no mechanism in SMTP to indicate that mail transited to a certain domain should be protected by TLS. In HTTPS, HSTS allows a web server to indicate that all future connections for a specified period of time must use HTTPS. However, in mail, messages are relayed in cleartext if TLS cannot be negotiated. As we showed in Section 4.1, this has led to organizations corrupting the STARTTLS negotiation to force mail to be sent in the clear. Whether this is being done for legitimate or nefarious purposes, it illustrates that STARTTLS provides no protection against frequently occurring man-in-the-middle attacks.

Second, even when TLS is used, there is no robust way for a sender to verify the authenticity of a recipient mail server. Common MTAs validate that a server's certificate matches the destination domain's MX record, not the destination domain name itself. However, this still leaves the server open to impersonation unless the DNS responses are separately authenticated. As we showed in Section 4.2, certain entities are using this weakness to reroute the flow of messages.

One potential option for preventing MITM attacks is to create a mechanism similar to HTTP Public Key Pinning for SMTP. This would allow a mail server to indicate whether future connections should require TLS and specify a public key. Other protections being adopted for the HTTPS certificate ecosystem might also be considered for STARTTLS, such as the use of Certificate Transparency [1] to guard against dishonest or compromised certificate authorities.

Finally, we note that end-to-end mail encryption, as provided by PGP [3] and S/MIME [31], does not address many of the challenges we discuss in this work. While these solutions do safeguard message content, they leave metadata, such as the subject, sender, and recipient, visible everywhere along the message's path. This information is potentially exposed to network-based attackers due to the lack of robust confidentiality protections for SMTP message transport. Although greater adoption of end-to-end encryption would undoubtedly be beneficial for security, for now, the overwhelming majority of messages depend solely on SMTP and its extensions for protection.

## 6.2 Challenges for Integrity

A major open question surrounding mail integrity is how to authenticate mail sent through mailing lists. Mailing lists frequently modify messages in transit and DKIM signatures are invalidated by these modifications, which prevents large mail providers from publishing a DMARC reject policy. When Yahoo deployed a reject policy in 2014, it resulted in a the heavy number of complaints and service malfunctions [8].

***[TODO: I'm not sure these points are correct. Neither of these are problems with the protocols themselves, but rather in how they are deployed. Re DKIM, couldn't you just pull that key from your DNS serevrs and the problem would go away?]*** A second challenge is ensuring strong integrity as organizations move to cloud providers, where mail infrastructure, IP address blocks, and machines, may be shared with other organizations. This infrastructure sharing is challenging in two respects. First, SPF has become less relevant, since, as explained in Section 5.1, SPF records tend to be overly broad. Second, DKIM becomes threatened by massive key compromises, as was the case for the SendGrid leak [4]. Overall, these two issues are part of a larger open question: How do we reliably establish the legitimacy of senders—whether for spam prevention or for integrity purposes—when many senders, good and bad, share common infrastructure?

The issue of shared infrastructure also affects mail confidentiality, as third-party providers would need certificates containing their clients' domains in order to enforce strict certificate verification. This is problematic, as it opens the door to attacks where the third-party mail provider—or an attacker who breaches their systems—uses these certificates to impersonate the clients' domains, either for mail delivery or for HTTPS connections. This threat might be mitigated with a scope-reducing X.509 extension or through some other mechanism not yet devised.

# 7. RELATED WORK

There has been little formal measurement of the public key infrastructure that supports mail transport. The most similar work is a set of Facebook blog posts that describe the STARTTLS configurations from the perspective of Facebook notifications [17, 18]. In May 2014, Facebook found that 28.6% of notification emails are transported over a STARTTLS connection with strict certificate validation, 28.1% are protected with opportunistic encryption (misconfigured STARTTLS server) and 41.0% of notifications are sent in cleartext. In August 2014, Facebook posted follow-up statistics, in which they note that 95% of notification emails are sent over STARTTLS with strict certificate validation. Facebook further notes that this rise is primarily due to two major mail providers, Yahoo and Microsoft, deploying STARTTLS . The jump of encrypted messages from 28.6% to 95% is incredibly exciting. However, as noted by Facebook, their notification emails are skewed towards personal addresses and large hosting providers, such as Gmail and Yahoo Mail.

In June 2014, Sean Rijs published a measurement study on the STARTTLS for 116 Dutch organizations, which found that 55% of tested domains used STARTTLS, 34% did not support STARTTLS, and 11% could not be tested [32]. Our results provide another perspective, including how incoming messages are protected, mail is authenticated, and organizations deploy STARTTLS . To the best of our knowledge, there have not been published studies on the deployment of DMARC, DKIM, SPF, nor evidence of widespread STARTTLS stripping or DNS servers lying

There is a large corpus of work on DNS servers providing false responses in order to facilitate content filtering [11, 26, 29, 35]. However, to the best of our knowledge, our study is the first to measure the amount the extent to which DNS servers are falsifying MX records for mail providers, and the amount of mail sent through these servers.

While Internet-wide scanning hasn't been used to measure the mail ecosystem, it has become a standard practice for measuring the HTTPS ecosystem. In 2010, the EFF performed a distributed scan of the IPv4 address space to identify certificate authorities. Later, in 2011, Holz et al. scanned the Alexa Top 1 Million in order to measure HTTPS deployment and commonly used certificate authorities [21]. In 2012, Heninger et al. performed comprehensive scans of the HTTPS to analyze the widespread use of weak cryptographic keys [19]. Again in 2013, Durumeric et al. completed daily scans in order to identify weaknesses in the HTTPS CA ecosystem [15].

In 2014, Huang et al. scanned the Top 1 Million to measure the deployment of Forward Secrecy [22].

# 8. CONCLUSION

While electronic mail carries some of users' most sensitive correspondence, SMTP did not originally include support for message confidentiality or integrity. Over the past fifteen years, the mail community has retrofitted SMTP with several security mechanisms, including STARTTLS, SPF, DKIM, and DMARC. In this work, we analyzed the global adoption of these technologies using data from two perspectives: Internet-wide scans that measured server configurations around the globe, and logs of SMTP connections to and from one of the world's largest mail providers over a sixteen month period. Our measurements show that use of these secure mail technologies has surged over the past year. However, adoption is not uniform, and many smaller organizations continue to lag in both deployment and proper configuration. The fail open nature of STARTTLS and the lack of strict certificate validation reflect the need for interoperability amidst the gradual rollout of secure mail transport, and they embody the old adage that "the mail must go through." Unfortunately, they also expose users to the potential for man-in-the-middle attacks, which we find to be so widespread that they affect more than 20% of messages delivered to Gmail from several countries. We hope that by drawing attention to these attacks and shedding light on the real-world challenges to secure mail, our findings will motivate and inform future research.

## Acknowledgements

***[TODO: add acknowledgements]*** The authors thank Eric Wustrow. We thank the exceptional sysadmins at the University of Michigan for their help and support. This material is based upon work supported by the National Science Foundation under grants CNS-1111699, CNS-1255153, CNS-1345254, CNS-1409505, CNS-1409758, and CNS-1518741, by the Google Ph.D. Fellowship in Computer Security, by the Morris Wellman Faculty Development Assistant Professorship, and by an Alfred P. Sloan Foundation Research Fellowship.

# 9. REFERENCES

[1] Certificate Transparency, 2015. http://www.certificate-transparency.org/.

[2] Alexa Internet, Inc. Alexa Top 1,000,000 Sites. http://s3.amazonaws.com/alexa-static/top-1m.csv.zip.

[3] J. Callas, L. Donnerhacke, H. Finney, D. Shaw, and R. Thayerj. OpenPGP message format, 2007. https://www.ietf.org/rfc/rfc4880.txt.

[4] D. Campbell. Update on Security Incident and Additional Security Measures, 2015. https://sendgrid.com/blog/update-on-security-incident-and-additional-security-measures/.

[5] Cisco. Cisco ASA 5500-X series next-generation firewalls, 2015. http://www.cisco.com/c/en/us/products/security/asa-5500-series-next-generation-firewalls/index.html.

[6] Cisco. Cisco IOS Firewall, 2015. http://www.cisco.com/c/en/us/products/security/ios-firewall/index.html.

[7] Cisco. SMTP and ESMTP inspection overview, 2015. http://www.cisco.com/c/en/us/td/docs/security/asa/asa93/configuration/firewall/asa-firewall-cli/inspect-basic.html#pgfId-2490137.

[8] L. Constantin. Yahoo email anti-spoofing policy breaks mailing lists, 2014. http://www.pcworld.com/article/2141120/yahoo-email-antispoofing-policy-breaks-mailing-lists.html.

[9] D. Crocker, T. Hansen, and M. Kucherawy. RFC 6376: Domainkeys identified mail (DKIM) signatures, Sept. 2011. https://tools.ietf.org/html/rfc6376.

[10] D. Crocker and T. Zink. M3AAWG trust in email begins with authentication, 2015. https://www.m3aawg.org/sites/maawg/files/news/M3AAWG_Email_Authentication_Update-2015.pdf.

[11] H. Duan, N. Weaver, Z. Zhao, M. Hu, J. Liang, J. Jiang, K. Li, and V. Paxson. Hold-on: Protecting against on-path DNS poisoning. In *Proc. Workshop on Securing and Trusting Internet Names*, 2012.

[12] Dukhovni, V. and Hardaker, W. and Parsons. SMTP security via opportunistic DANE TLS, July 2013. http://tools.ietf.org/html/draft-ietf-dane-smtp-with-dane-12.

[13] Z. Durumeric, D. Adrian, J. Kasten, D. Springall, M. Bailey, and J. A. Halderman. POODLE Attack and SSLv3 Deployment, 2014. https://poodle.io.

[14] Z. Durumeric, M. Bailey, and J. A. Halderman. An Internet-wide view of Internet-wide scanning. In *Proc. 23rd USENIX Security Symposium*, 2014.

[15] Z. Durumeric, J. Kasten, M. Bailey, and J. A. Halderman. Analysis of the HTTPS certificate ecosystem. In *Proceedings of the 13th Internet Measurement Conference*, Oct. 2013.

[16] Z. Durumeric, E. Wustrow, and J. A. Halderman. ZMap: Fast Internet-wide scanning and its security applications. In *Proceedings of the 22nd USENIX Security Symposium*, Aug. 2013.

[17] Facebook. The current state of SMTP STARTTLS deployment, May 2014. https://www.facebook.com/notes/protect-the-graph/the-current-state-of-smtp-starttls-deployment/1453015901605223/.

[18] Facebook. Massive growth in SMTP STARTTLS deployment, Aug. 2014. https://www.facebook.com/notes/protect-the-graph/massive-growth-in-smtp-starttls-deployment/1491049534468526.

[19] N. Heninger, Z. Durumeric, E. Wustrow, and J. A. Halderman. Mining your Ps and Qs: Detection of widespread weak keys in network devices. In *21st USENIX Security Symposium*, Aug. 2012.

[20] P. Hoffman. RFC 3207: SMTP service extension for secure SMTP over transport layer security, Feb. 2002. http://www.ietf.org/rfc/rfc3207.txt.

[21] R. Holz, L. Braun, N. Kammenhuber, and G. Carle. The SSL landscape: A thorough analysis of the X.509 PKI using active and passive measurements. In *11th ACM SIGCOMM conference on Internet measurement (IMC)*, 2011.

[22] L.-S. Huang, S. Adhikarla, D. Boneh, and C. Jackson. An experimental study of TLS forward secrecy deployments. In *Web 2.0 Security and Privacy (W2SP)*, 2014.

[23] S. Kitterman. RFC 7208: Sender policy framework (SPF) for authorizing use of domains in email, Apr. 2014. http://tools.ietf.org/html/rfc7208.

[24] J. Klensin. RFC 5321: Simple mail transfer protocol, Oct. 2008. http://tools.ietf.org/html/rfc5321.

[25] M. Kucherawy and E. Zwicky. RFC 7489: Domain-based message authentication, reporting, and conformance (DMARC), Mar. 2015. https://tools.ietf.org/html/rfc7489.

[26] G. Lowe, P. Winters, and M. L. Marcus. The great DNS wall of China. *New York University*, 21, 2007.

[27] Microsoft. Tls functionality and related terminology, June 2014. http://technet.microsoft.com/en-us/library/bb430753%28v=exchg.150%29.aspx.

[28] Mozilla Developer Network. Mozilla network security services (NSS). http://www.mozilla.org/projects/security/pki/nss/.

[29] Z. Nabi. The anatomy of web censorship in pakistan. *arXiv preprint arXiv:1307.1144*, 2013.

[30] J. B. Postel. RFC 821: Simple mail transfer protocol, Aug. 1982. http://tools.ietf.org/html/rfc821.

[31] B. Ramsdell and S. Turner. Secure/multipurpose Internet mail extensions (S/MIME) version 3.2 message specification, 2010. https://tools.ietf.org/html/rfc5751.

[32] S. Rijs and M. van der Meer. The state of StartTLS, June 2014. https://caldav.os3.nl/_media/2013-2014/courses/ot/magiel_sean2.pdf.

[33] smtp_tls_security_level. http://www.postfix.org/postconf.5.html#smtp_tls_security_level.

[34] Verisign Labs. DNSSEC scoreboard, 2015. http://scoreboard.verisignlabs.com/.

[35] J.-P. Verkamp and M. Gupta. Inferring mechanics of web censorship around the world. *Proc. 3rd USENIX Free and Open Communications on the Internet (FOCI)*, 2012.