The Robustness of New Email Identification Standards

Patrik Ostrihon, ComDom Software, patrik.ostrihon@comdomsoft.com

Reza Rajabiun, ComDom Software and York U. reza.rajabiun@comdomsoft.com

The Problem

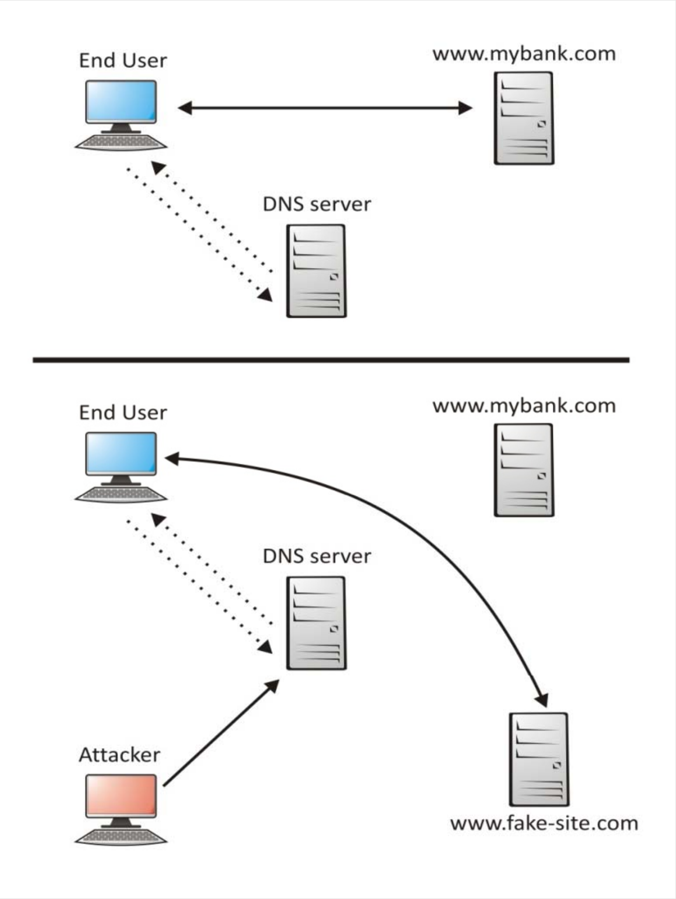
- The Spam Puzzle: Growth in level and sophistication of Spam, despite increased filter accuracy.
- Multilayer Filtering or the Dangerous Econ. of Spam Control (Kimakova and Rajabiun, 2008 MIT Spam Conference.)
- This paper focuses on a specific and small subset of mechanism enhancements.

COM DOM ANTISPAM

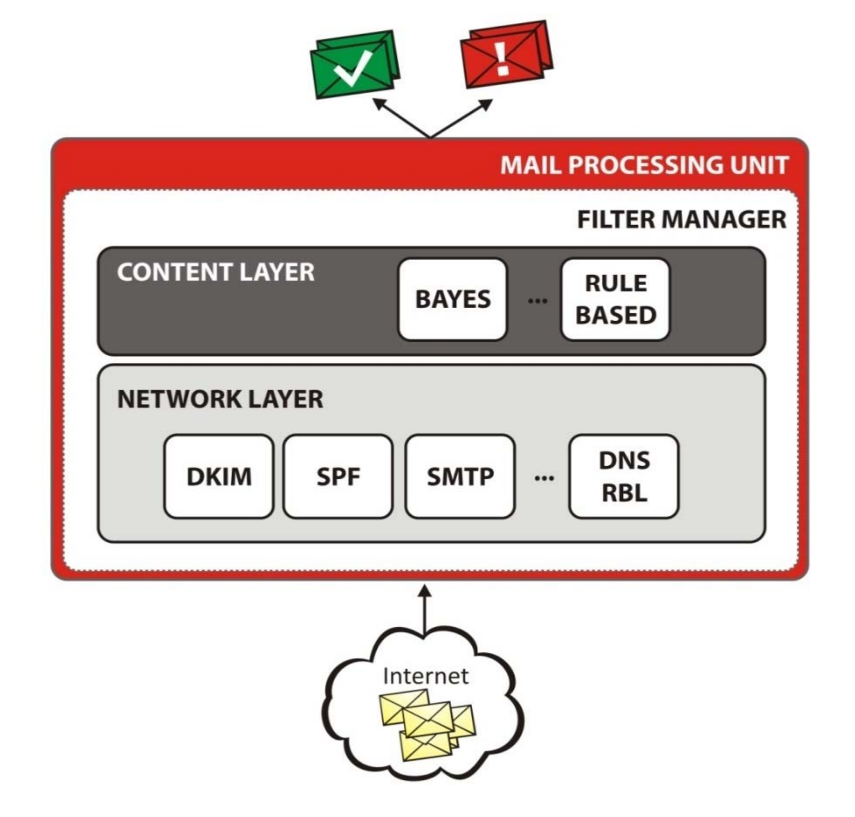
New Identification Standards

- Semantic note: Authentication versus Identification.
- Important link between authentication/identification, and functioning of reputation systems
- The robustness of DKIM and SPF, as representative of different classes of similar mechanisms
- Objective of both mechanisms: Limit abuse of well known vulnerabilities of SMTP and DNS (DNS Poisoning)
- Research question: Complements or substitutes to statistical content filters?

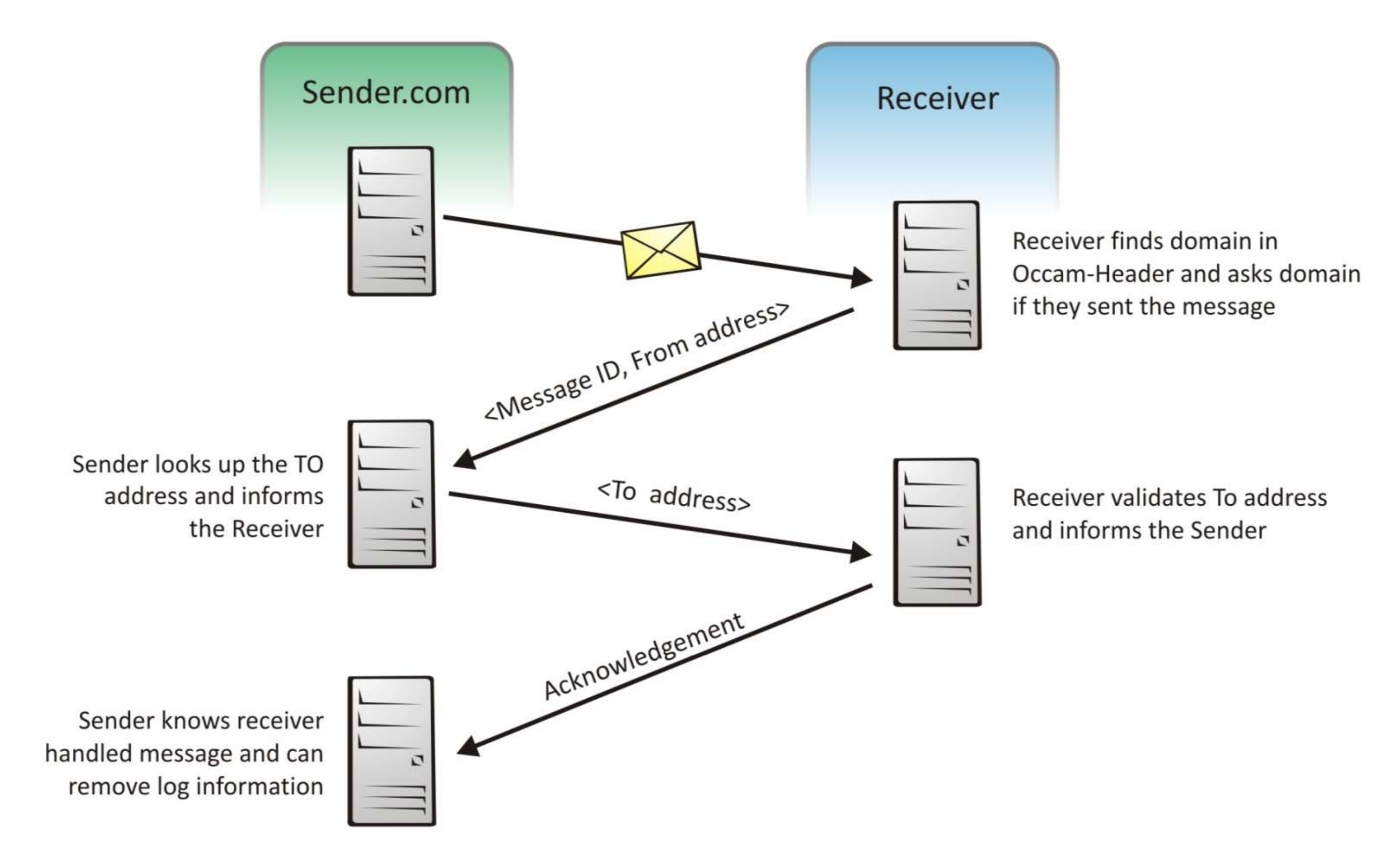
DNS Spoofing



Typical Multilayer Filter



Receiver (DKIM/SPF) vs. Sender Side Auth. (Fleizach et al. (2007)



Receiver finds domain in Occam-Header and asks domain if they sent the message

Sender looks up the TO address and informs the Receiver

Receiver validates To address and informs the Sender

Sender knows receiver handled message and can remove log information

SPF/DKIM

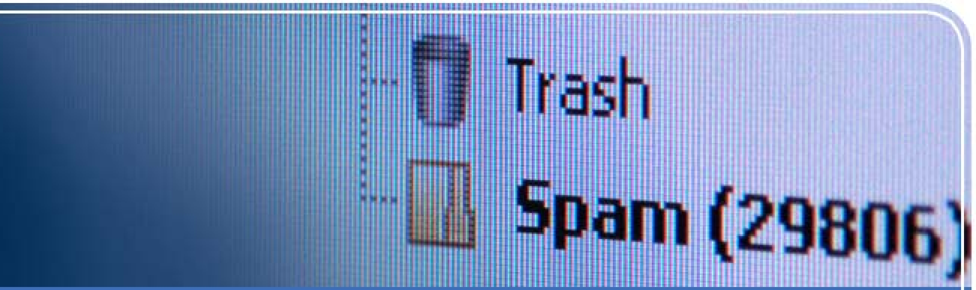
- Defined in: IETF RFC 4871 and RFC 4408
- Impose burden of proof of the identity is valid/not on receivers (fixed and variable costs of enhancement)
- Limited data on adoption (SPF: app. 15%, DKIM: Bulk mailers/large ISPs)
- Why? Ease of subversion or switching costs?
- Senders: Adopting all, lower false positives

COM DOM ANTISPAM

SMTP/DNS vul. to be addressed

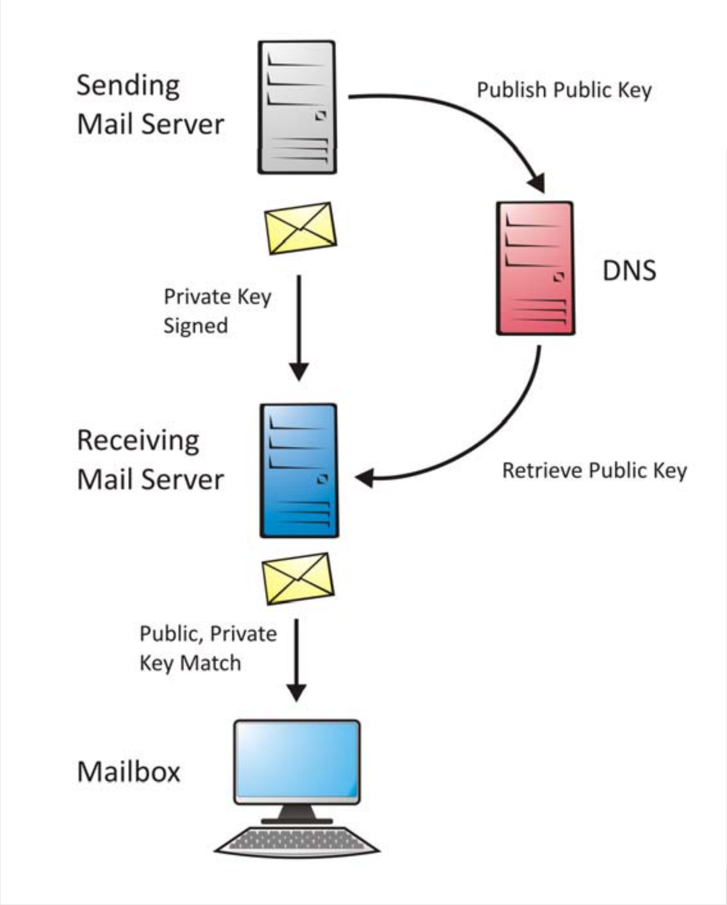
- Ozment and Schechter (2006)
- 1) DDOS: Making a system unavailable to users
- 2) Man in the middle problem: Interception of com. between clients and hosts, forge identities and content
- 3) Compromised servers: Alter integrity of DNS records before requested by client

DKIM



- Senders or intermediaries cryptographically sign messages.
- First Q: How many signatures?
- Receivers query DNS servers of senders for public key.
- In practice MTA insert sign. in transit
- Chain of trust among semi-autonomous nets of large ISPs and senders of bulk emails

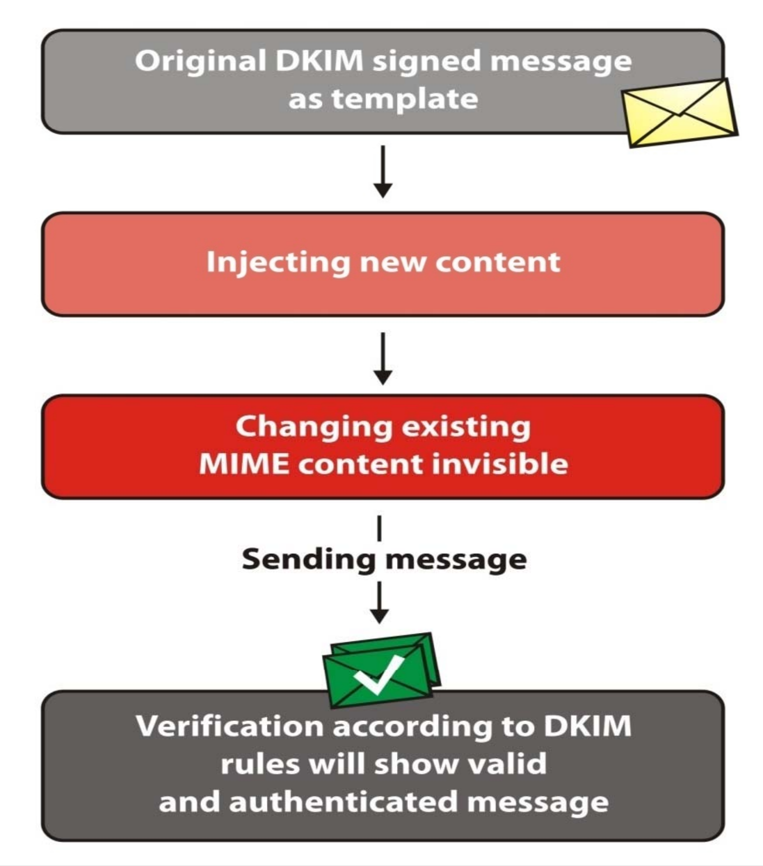
DKIM Architecture



DKIM Problems

Fundamental separation of sending/signing authority

- Entity that signs a message also authority to define domain name later used by receiver to assess message quality
- State of Spam tech: Easy to infiltrate servers and copy signatures of large ISPs.
- + One shot BGP Spectrum Agility tech.
- + Delay, comp/com burden (2.5x increased latency, Fleizach et al. 2007)



Original DKIM signed message as template

Injecting new content

Changing existing MIME content invisible

Sending message

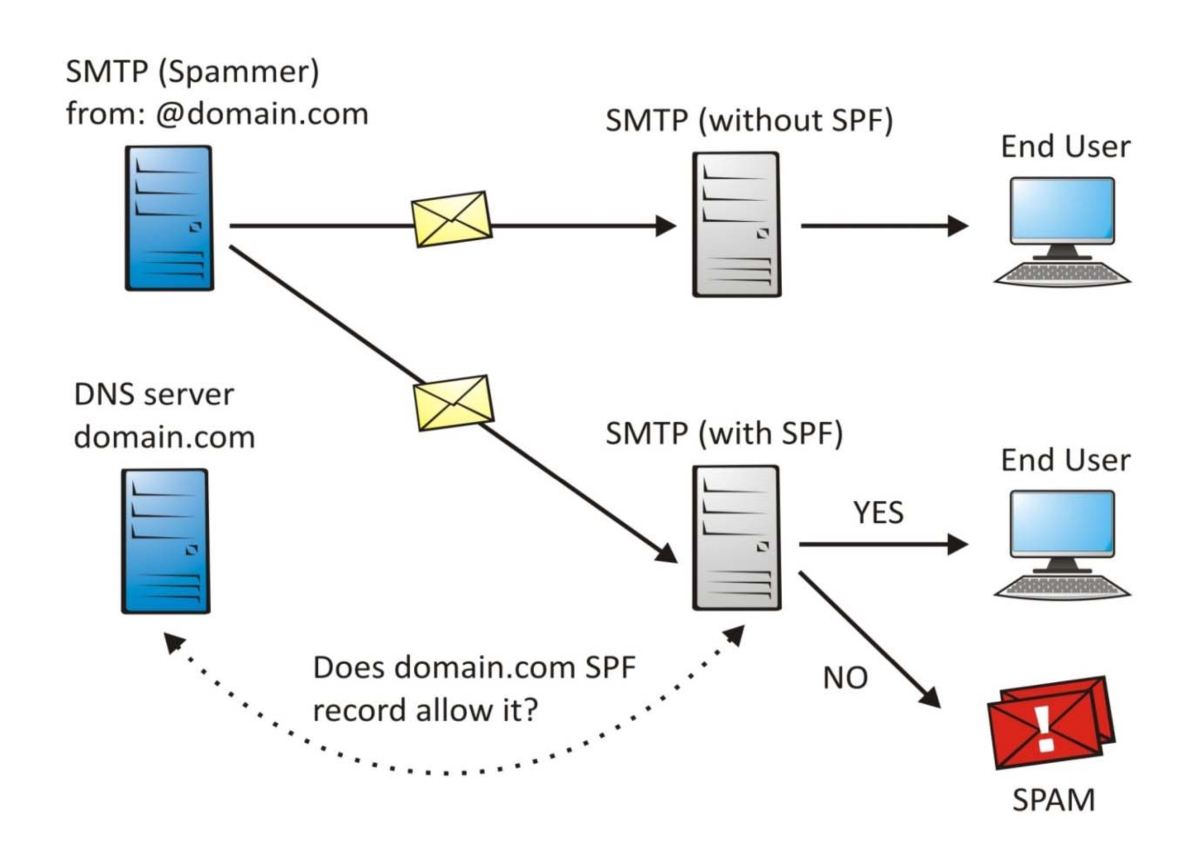
Verification according to DKIM rules will show valid and authenticated message

COM DOM ANTISPAM

SPF

- An extension of SMTP.
- Allows software to identify and reject forged addresses in the SMTP Mail From (Return-Path)
- MAAWG (2008): As “path registration” (vs. authentication).
- Generally: Providing domain owners with a set of rules for who (which host in that domain) is authorized to send (sender origin)
- As DKIM: DNS Poisoning?

Trash
Spam (29806)


SMTP (Spammer)
from: @domain.com
SMTP (without SPF)
End User
DNS server
domain.com
SMTP (with SPF)
End User
YES
NO
Does domain.com SPF
record allow it?
SPAM

SPF Architecture

- Rules of authorization from very simple (IP address listing) to very complex
- Principles of operation: Rule definitions implemented via DNS's TXT record (similar to DNSBL)
- Except: SPF exploits authority delegation scheme of real DNS
- DNS queries cached on server side

SPF Process

- Can lower error messages/auto-reply (back scatter)
- SPF allows: users to identify their legitimate sending IP with a FAIL result for all other Ips.
- Receivers then can check SPF records and reject forgeries
- Benefit: Mainly to senders whose email addresses are forged in the Return-Path.

SPF Problems

Multifaceted:

a) Messages that go through intermediaries (forwarding, hosting)

- ➔ Hence: Increasing prob. of false positives
- ➔ This problem can be easily fixed by: 1) replacing the original sender with one belonging to the local domain, 2) refusing (answering 551user not local, please try user@example.com), 3) Sender Rewriting Scheme

SPF Problems Cont.

b) Persistence of compromised systems on domains that take advantage of SPF

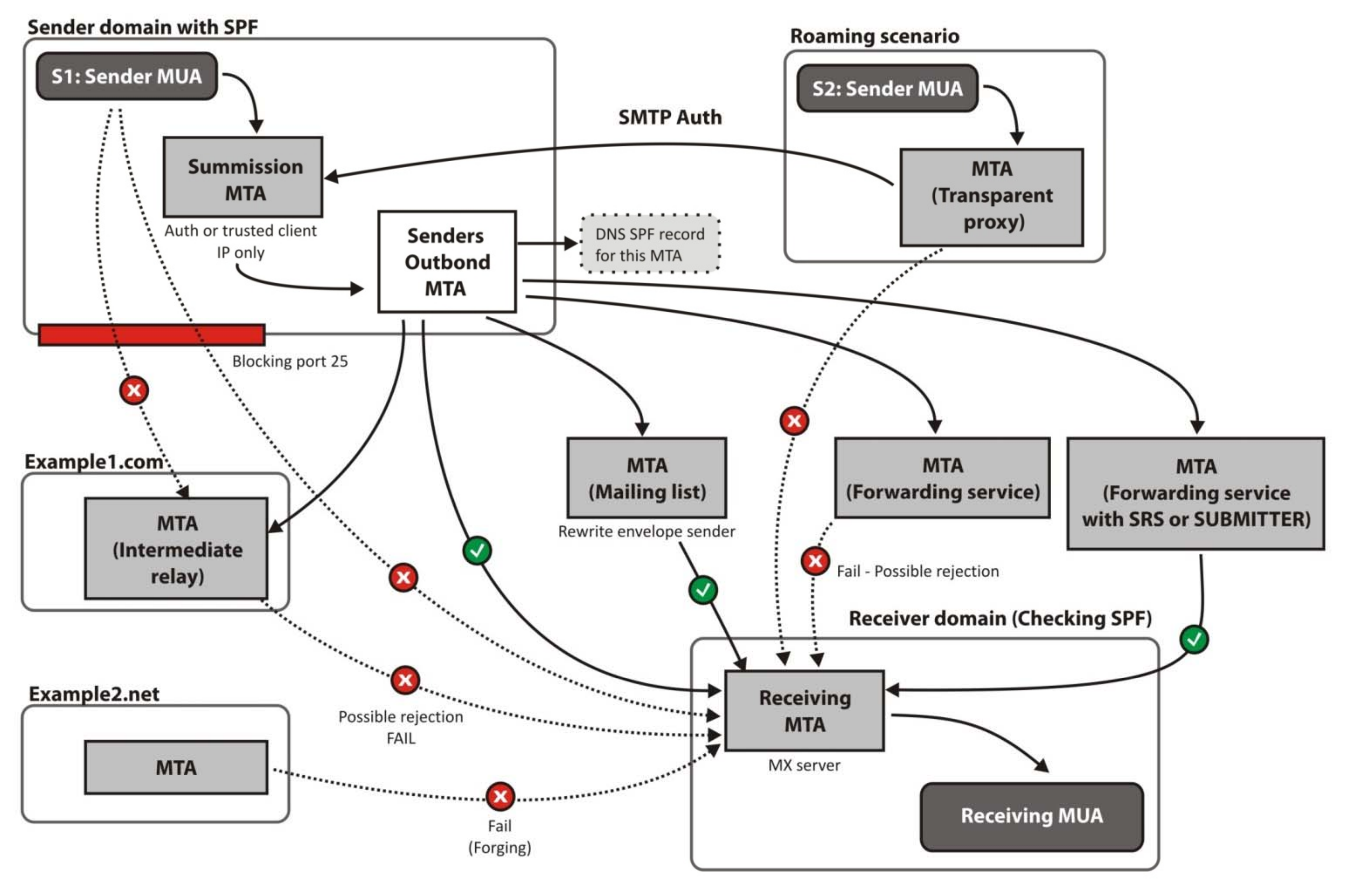
c) Can be used as an instrument of DoS (2006 IETF draft)-response by SPF Project

- → Limit of 10 SPF mechanisms, each can generate 10 queries = 100 transactions for each name to be resolved
- → Also: Can use local macros to randomize further queries (where 0 spammer resources are used)
- → Infinite gain DNS amplification attack!

SPF Implementation Experience

- SPF can be useful, only when rules specified in DNS records are restrictive.
- Reasonable default policies (those that apply where there are no specific rules.
- Unhelpful policies: a) + all (PASS), b) ?all (SOFTFAIL), C) ~all (NEUTRAL)
- Only useful: -all (Fail): Because the only way to tell another mail server not to accept messages from unauthorized senders + minimize backscatter

SPF in the wild



SPF Other

a) Can help a little, but does NOT validate that a message comes from a claimed user. Users within one domain can forge each other's addresses. (big problem for large ISPs)

b) Difficulties in interpreting SOFTFAIL (news letters, bills....) Why email marketeers don't like SPF, and prefer DKIM.

c) Checking SPF behind “border MTA”, possible, but too late to reject SPF FAIL. Can only delete FAILing mail

d) High DNS amplification attack/Spammers resources

Implications

- Neither very robust to current spamming technologies
- DKIM to server hacking and man in the middle problems: Used to build a chain of trust between large commercial senders and network operators
- SPF: Lower resource footprint, backscatter, but the risk of attacks and increased risk of false positives
- In the broader multilayer filtering context: Marginal value of information from the two not very high.
- Q for discussion: Identification (authentication/reputation) enhancements, content filters: Complements or Substitutes in 5-10 years?

The Robustness of New Email Identification Standards

Patrik Ostrihon, ComDom Software, patrik.ostrihon@comdomsoft.com

Reza Rajabiun, ComDom Software and York U. reza.rajabiun@comdomsoft.com

COM DOM ANTISPAM