

Spamszűrési módszerek és eljárások

Kadlecsik József

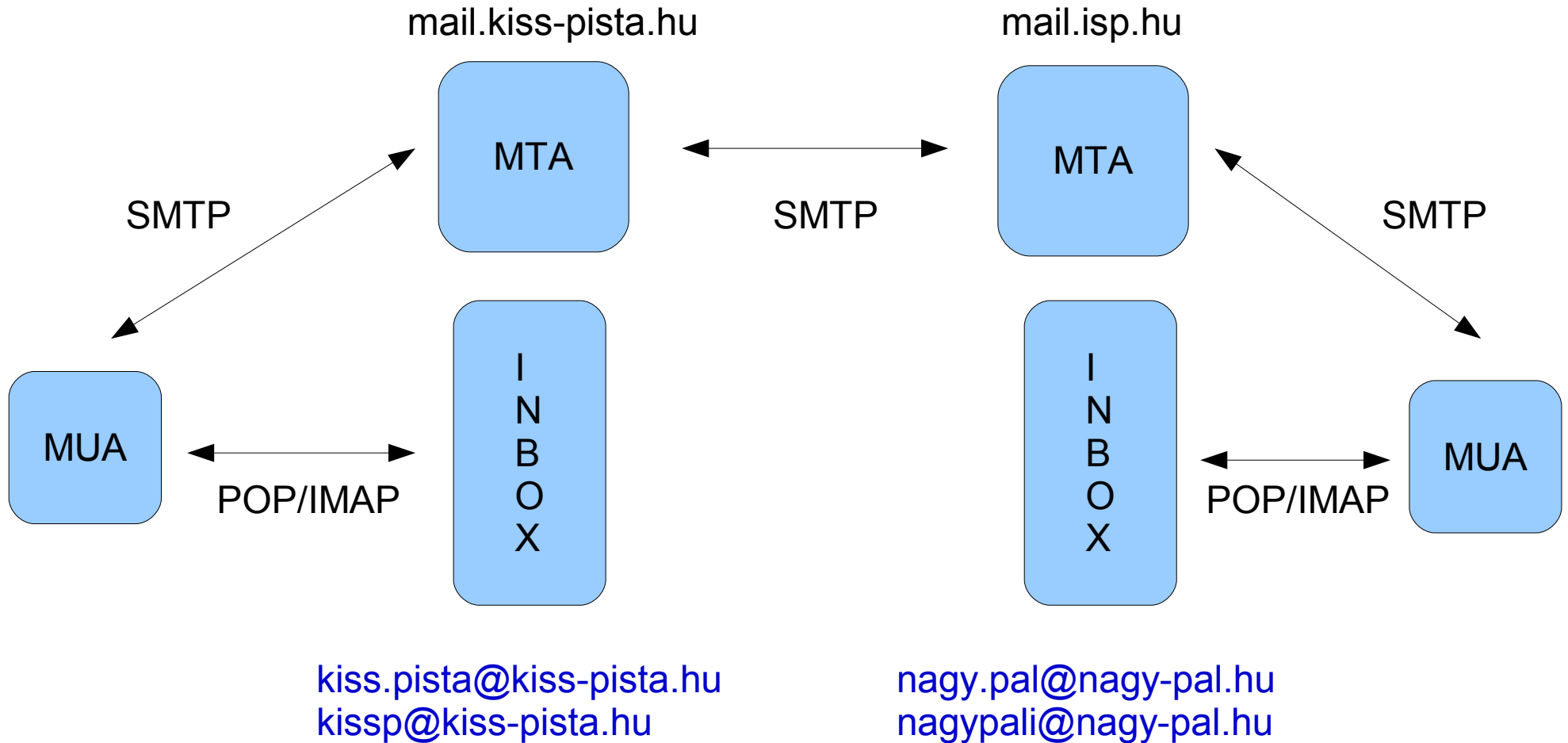
KFKI RMKI

<kadlec@sunserv.kfki.hu>

Tartalom

- Az elektronikus levelezés működés
- Spammer technikák
- Védekezési- és spamszűrési módszerek

Az elektronikus levelezés működése



Az SMTP és a DNS

- @domain-part: MX vagy A rekord
- MX rekord: súly és A rekord (**nem** CNAME):

```
% dig -t mx domain-part
```
- Backup MX rekord
- Wildcard A/MX rekord
 - 2003.09.15: Verisign SiteFinder: .com. .net
 - hibadetektálási nehézségek
 - potenciális E-mail cím gyűjtés
 - hamisított (nem létező) E-mail címek valódinak tűnnek

SMTP session

```
% telnet test.szerver.hu smtp
220 test.szerver.hu ESMTP Postfix
EHLO test.kliens.hu
250-test.szerver.hu
250-PIPELINING
250-SIZE 10240000
250-ETRN
250 8BITMIME
```

SMTP session folyt.

MAIL FROM: <teszt@test.kliens.hu>

250 Ok

RCPT TO: <teszt@teszt.szerver.hu>

250 Ok

DATA

354 End data with <CR><LF>.<CR><LF>

.

250 Ok: queued as A342A1F1300

QUIT

SMTP és E-mail fejlődés

- SMTP:

```
MAIL FROM: <joe@somewhere.com>
```

```
RCPT TO: <target@mail.host>
```

- E-mail fejlődés:

```
From: George.W.Bush@whitehouse.gov
```

```
To: Condoleezza.Rice@whitehouse.gov
```

SMTP sarokpontok

- Levél (nyomtalanul) nem vehet el:
 - bounce
 - double-bounce
- A '.'-ra válaszul adott '250 Ok' üzenettel a levél kezelésének gondját a fogadó szerver veszi át.
- A '.'-ra a válasz nem jöhet “túl soká”:
 - levél-duplikálás

Spam formák

- Kéretlen E-mail
 - illegális, fél-legális kereskedelem
 - pornó
 - pénzügyi beugratások, csalás
 - phishing
- Vírusos üzenetek
- Joe-jobbing
- Backscatter

Spammerek céljai

- Nagy tömegű E-mail küldése:
 - open relay-ek
 - lyukas CGI/PHP script-ek (formmail)
 - open proxy (zombi) gépek – vírusok
 - offshore ISP-k: Kína, Dél-Korea, Indonézia, Malajzia, volt Szovjet államok, Dél-Amerika, stb.
 - *pink contract*: neves ISP mögött felárért

Spammerek céljai, folyt. I.

- Rejtőzködés, nyomok elfedése
 - host ISP mail szerverének elkerülése: direct-to-MX szoftverek
 - hamisított E-mail fejlécek (Received, Message-ID, stb.)
 - hamisított To: és From: E-mail fejlécek

Spammer trükkök

- Tartalomszűrők kikerülése:
 - Kódolt/valódi mögé rejtett URL-e; MIME, URL, HTML kódolások alkalmazása
 - Random karakterek, szövegrészek beillesztése: böngésző a felhasználó felé nem (észrevehetően) mutatja (MIME, HTML, CSS trükkök), tartalomszűrőt félrevezeti
 - Szándékosan hibás írásmód
 - Hibás, nem létező HTML tag-ek

Példák

Spammer trükkök, folyt.

- DNS játékok:
 - eldobható DNS nevek spam küldésére; domain kiting
 - különböző IP blokkból több IP cím a webserverek; portálok és társcégek
 - az IP címek (akár több tucatnyi) gyors rotálása
 - DNS kikapcsolása, miután a DNS cache-k megtanulhatták
 - DNS válasz késleltetése/letiltása az automatizált spamkeresők számára (SpamCop)
 - wildcard DNS rekordok

Védekezési lehetőségek

- SMTP szintű védelem:
 - before-queue
- Tartalomszűrés
 - before-queue: potenciális problémák
 - after-queue

SMTP szintű védelem I.

- SMTP szintaktika és protokoll megkövetelése:
 - HELO/EHLO szükséges
 - Érvényes HELO/EHLO név: szintaktikailag korrekt és FQDN
 - Érvényes MAIL FROM/RCPT TO szintaxis
 - ESMTP pipelining betartatása
- Előnyök és hátrányok:
 - egyszerű alkalmazni
 - rosszul megírt home-breed szoftverek, partner-MTA-k
 - SMTP-t beszélő spammer szoftverek terjedés

SMTP szintű védelem II.

- Feladó és címzett domain létezzék: MX/A
 - bounce küldés elvileg lehetséges
 - feladó (domain) hamisítása
 - korrekt DNS rekordok, wildcard rekord
- Feladó E-mail cím létezésének ellenőrzése:
 - nemlétező feladói címek kiszűrése
 - feladó E-mail cím hamisítása
 - timeout problémák
 - nézhetik DHA támadásnak
 - intermediate relay hostoknál nem működik
 - Yahoo és társainál nem működik (DATA vs RCPT)

SMTP szintű védelem III.

- Címzett E-mail cím létezésének ellenőrzése:
 - nemlétező címzettek kiszűrése: bounce nem a mi dolgunk
 - DHA-val az érvényes E-mail címek begyűjthetők
- Multi-recipient bounce elutasítása

SMTP szintű védelem IV.

- Kliens IP cím-név DNS feloldásának ellenőrzése:
 - korrekt DNS az Internet sarokköve; zombi gépek ellen jó
 - sok false positive
- HELO/EHLO név DNS feloldásának ellenőrzése
 - korrekt DNS az Internet sarokköve; zombi gépek ellen jó
 - még több false positive
- HELO/EHLO név, feladó/címzett domain NS/MX rekordjának az ellenőrzése
 - általában inkább kivételek felállítására
 - hamisítás nem kivédhető

SMTP szintű védelem V.

- RBL/RHSBL listák
 - disztributált, akár igen gyors védelem
 - listák minősége változó
 - spammerek sikeresen támadtak listagazdákat:
 - Osirusoft, monkeys.com, blackhole.compu.net, ...
- RBL/RHSBL lista típusok:
 - spammer források, spammer csapdák, automatizált listák
 - viselkedés-alapú, tartalom (URL)
 - zombi hálók
 - open relay, proxy, formmail
 - dinamikus IP tartományok, ország-blokkok

SMTP szintű védelem VI.

- Greylisting: kliens IP cím, feladó és címzett E-mail cím hármásának ellenőrzése: láttuk-e már?
 - igen hatékony a nem MTA spammer/vírus szoftverek ellen
 - egyedi id E-mail címeknél, szerver-farm esetén nem működik (kivételek)

SMTP szintű védelem VII.

- Számlálók használata: kliens IP címe + ismeretlen címzettek száma egy periódus alatt (DHA)
- Számlálók használata: HELO/EHLO randomizáció
 - viszonylag egyszerű
 - sender verification próbák, szerver-farmok
- Hamisítások kiszűrése: HELO/EHLO név
 - egyértelműen spammer
 - roaming user
- Spammer csapdák

SMTP szintű védelem VIII.

- Challenge-response rendszerek
 - egyértelmű kizárása a spam-nak
 - udvariatlan, nem szokványos
- Felhasználónként címlisták
 - egyértelmű kizárása a spam-nak
 - frissítési, karbantartási kérdések

Tartalomszűrés I.

- Egyszerű regexp/pcre mintakeresés header/body-ban
 - backscatter és tartalomszűrés
 - elrejtések
 - karbantartás, teljesítmény problémák
- Nyelv és kódolás
 - könnyű
 - előre nem látható partnerek

Tartalomszűrés II.

- Bayesian szűrők
 - nagy választék: spamassassin, bogofilter, dspam, stb.
 - mi a ham és mi a spam?
 - memória és diszk igény
- OCR
- Elméleti háttér

Tartalomszűrés III.

- Elosztott szűrők: Vipul's Razor, DCC, stb.
 - disztributált, gyors reagálás
 - random szavak, szövegrészek

Kiegészítő módszerek I.

- Sender Policy Framework (SPF)
- E-mail címtartomány hozzákötése az engedélyezett SMTP szerverekhez
- SMTP szinten működik:
 - forward
 - mail relay
 - SRS (Sender Rewriting Scheme) nem segít

Kiegészítő módszerek II.

- DomainKeys (DKIM)
- E-mail autentikáció a DomainKey-Signature fejlécen és DNS bejegyzéseken keresztül
 - message body: levlisták

Összegzés helyett

- Nem állunk vereségre
- Nem állunk győzelemre