# Aktivní ochrana proti spamu

Active Anti-spam

Bc. Jan Havlíček

Diplomová práce
2007



Univerzita Tomáše Bati ve Zlíně
Fakulta aplikované informatiky

Univerzita Tomáše Bati ve Zlíně
Fakulta aplikované informatiky
Ústav aplikované informatiky
akademický rok: 2006/2007

# ZADÁNÍ DIPLOMOVÉ PRÁCE
(PROJEKTU, UMĚLECKÉHO DÍLA, UMĚLECKÉHO VÝKONU)

| | |
|---|---|
| Jméno a příjmení: | **Bc. Jan HAVLÍČEK** |
| Studijní program: | **N 3902 Inženýrská informatika** |
| Studijní obor: | **Informační technologie** |
| Téma práce: | **Aktivní metody řešení problematiky spammingu** |

Zásady pro vypracování:

**Předmětem řešení práce je návrh a implementace bezpečnostního řešení spammingu. Práce je řešena formou projektu který analyzuje rizika a stanovuje vhodné postupy řešení.**

**1. Analyzujte problematiku internetového spammingu v různých informačních kanálech (email, instant messaging, internetová telefonie).**

**2. Porovnejte náročnost, efektivnost a bezpečnost jednotlivých řešení ochrany proti spammingu**

**3. Navrhněte a realizujte projekt řešení antispamového zabezpečení emailového serveru s důrazem na metody nevyžadující analýzu obsahu zprávy**

**4. Zhodnoťte pozitiva a negativa navrženého řešení a stanovte jeho přínosy.**

Rozsah práce:

Rozsah příloh:

Forma zpracování diplomové práce: **tištěná/elektronická**

Seznam odborné literatury:

**Srinivasa. RADER, Devin.ASP.NET 2.0 Programujeme profesionálně: Computer Press, 2007**
**DUTHIE, G. Andrew. MS ASP.NET Krok za krokem: Computer Press, 2005**
**AVERY, James. Microsoft ASP.NET Konfigurace a nastavení: Computer Press, 2004**
**LACKO, Luboslav. ASP.NET a ADOSP.NET 2.0: Computer Press, 2006**

Vedoucí diplomové práce: **doc. Mgr. Roman Jašek, Ph.D.**
Ústav informatiky a statistiky

Datum zadání diplomové práce: **13. února 2007**

Termín odevzdání diplomové práce: **28. května 2007**

Ve Zlíně dne 13. února 2007

prof. Ing. Vladimír Vašek, CSc.
*děkan*

L.S.

UNIVERZITA TOMÁŠE BATI VE ZLÍNĚ
Fakulta aplikované informatiky

doc. Ing. Ivan Zelinka, Ph.D.
*ředitel ústavu*

## ABSTRAKT

Tato práce se zabývá problematikou spamu zejména v oblasti, kde je tento problém nejrozšířenější, tedy prostřednictvím elektronické pošty. Určuje, co je to spam, a shrnuje možné způsoby ochrany proti němu. V případě elektronické pošty obsahuje stručný popis způsobu doručování zpráv prostřednictvím protokolu SMTP a popisuje metody ochrany proti spamu z pohledu koncového uživatele a jeho chování, dále z pohledu infrastruktury přijímající emaily a také z pohledu infrastruktury emaily odesílající. Okrajově se též dotýká očekávané problematiky spamingu prostřednictvím diskusních fór, VoIP (Voice over IP – Internet telephony) a IM (instant messagingu). V praktické části je navrženo konkrétní řešení antispamové ochrany serveru elektronické pošty.

Klíčová slova: email, SMTP, spam, instant messaging, VoIP, greylisting, SPF, postfix

## ABSTRACT

This thesis deals with questions of spam in electronic mail, which is the field where spam is most prominent. The work specifies what spam is and summarizes possible methods of anti-spam protection. A brief description of message delivery through SMTP protocol is provided in the section dedicated to electronic mail, in further parts of the thesis the methods of anti-spam protection are discussed. These methods concern the end-users' point of view and their behavior, the view of infrastructure receiving emails, and the view of infrastructure dispatching the mail. Anticipated problems of spam in discussion groups and through VoIP (Voice over IP – Internet telephony) and IM (Instant Messaging) are mentioned as well. The practical part of the thesis offers a concrete solution of an e-mail server anti-spam protection.

Keywords: email, SMTP, spam, instant messaging, VoIP, greylisting, postfix

Rád bych na tomto místě poděkoval vedoucímu práce, Doc. Romanu Jaškovi, za ochotu ujmout se vedení práce a cenné konzultace v průběhu jejího zpracování.

Dále bych vyslovil poděkování Mgr. Zuzaně Fenclové za podporu při psaní práce a pomoc s její jazykovou úpravou. Dík též patří občanskému sdružení PinkNET.cz za poskytnutí záznamů provozu emailového serveru pro zhodnocení účinnosti vybraných antispamových metod.

Prohlašuji, že jsem na diplomové práci pracoval samostatně a použitou literaturu jsem citoval. V případě publikace výsledků, je-li to uvolněno na základě licenční smlouvy, budu uveden jako spoluautor.

Ve Zlíně …….

Podpis diplomanta

# OBSAH

# ÚVOD

V následujícím textu se budeme zabývat elektronickou výměnou zpráv prostřednictvím různých používaných kanálů a zvláštním fenoménem s tím souvisejícím, tedy zahlcováním koncových uživatelů (čtenářů) tzv. spamem.

Přesná definice pojmu spam je nesnadná, neboť posouzení je svým způsobem subjektivním problémem každého čtenáře. Rozumnou definici nám nabízí např. Encyklopedie Wikipedia CS [1], a to takovouto: „Spam je nevyžádané masově šířené sdělení (nejčastěji reklamní) šířené internetem.“ Subjektivně příjemce pošty obvykle rozhoduje dle dvou kritérií:

- zda autor své sdělení cílil na adresáta, kterému byla zpráva doručena a který ji čte jako na konkrétní osobu nebo jako na příslušníka nějaké přiměřeně rozumně definované skupiny
- zda se čtenář cítí čtením (případně jiným nakládáním se sdělením, jako například odstraněním) obtěžován a omezován.

Za spam tedy obvykle označujeme sdělení doručené nějakému konkrétnímu adresátovi bez alespoň přiměřeného cílení odesílatele na tohoto konkrétního adresáta, sdělení obvykle mající charakter obchodní zprávy. Subjektivita rozhodování spočívá v tom, že až čtenář zprávy rozhoduje, zda je pro něj sdělení zajímavé, či obtěžující (dámy zřejmě nebudou považovat za příliš zajímavou nabídku pilulek na prodloužení penisu, profesor vysoké školy zase nabídky na zakoupení bakalářského diplomu nějaké světově proslulé školy v Grónsku). Sdělení nemusí nutně být obchodní či komerční, obdobným způsobem mohou být šířena sdělení náboženská, politická, ideologická, případně i zprávy svým obsahem již spadající pod aktivity vysloveně kriminální (phishing, pharming aj.). Technicky v podstatě stejným způsobem se šíří i e-mailové viry, které se obvykle rozesílají na veškeré adresy nalezené na již napadeném počítači. Doplňujícím předpokladem, i když z pohledu postiženého čtenáře ne tak důležitým, je skutečnost, že sdělení bylo distribuováno masově, tedy většímu množství adresátů.

Se spamem se setkáváme svým způsobem i mimo elektronickou komunikaci; až na výjimku podmínky šíření pomocí internetu se pod spaming dá zařadit každodenní zahlcování běžné poštovní schránky přídělem letáků. Fenomén spamingu nicméně podnítilo právě elektronické prostředí, které umožnilo doručovat s velmi malými náklady

velké množství zpráv a vzhledem k nepřipravenosti komunikačních protokolů i možnost skrývání identity odesilatele (obtěžovatele).

Cílem této práce je zejména definice problematiky spamingu, vyhledání, popis a analýza metod boje proti spamu a v praktické části návrh organizačních i technických antispamových opatření použitelných provozovatelem rozsáhlejší komunikační infrastruktury, kterou lze považovat za organizačně uzavřenou, jako je například větší firma, síť university a podobně. V práci se budeme zabývat především spamem v elektronické poště, okrajově se zmíníme i o jiných v současné době známých a využívaných formách spamingu, zejména spamu komentářovém, Usenet news spamu, spamu v prostředí IM a spamu v prostředí VoIP a možných metodách ochrany proti nim.

# I. TEORETICKÁ ČÁST

# 1. ELEKTRONICKÁ POŠTA

Nejstarší metodou využívanou pro předávání zpráv mezi komunikujícími stranami je elektronická pošta. Vzhledem k tomu, že za univerzální přenosové medium pro transport elektronické pošty je považován Internet, bude předmětem zkoumání tohoto textu pouze elektronická pošta založená na protokolu SMTP tak, jak je v Internetu a IP sítích provozována. Ostatní v historii vzniklé systémy elektronické pošty jsou dnes buď nahrazeny přenosy založenými na SMTP (UUNET, …), nebo jsou provozovány jako proprietární, obvykle s možností propojovacího můstku do systému SMTP pošty (Lotus Notes, Exchange, …).

## 1.1. Protokol SMTP historie s způsob fungování

V podstatě souběžně se vznikem projektu ARPANET a rodiny protokolů TCP/IP vznikl i aplikační protokol SMTP – Simple Message Transport Protocol, jak je přímo z názvu patrno, protokol určený k předávání jednoduchých zpráv. Prvotní definice protokolu SMTP byla kodifikována v RFC 821 [11]. Stejně jako mnoho dalších komunikačních protokolů z této ranné doby byl SMTP v průběhu doby doplňován o další funkčnosti, avšak základní jednoduchost mu zůstala. V současné době je protokol používán v rozšířené podobě označované jako ESMTP a je kodifikován v RFC 2821 [12], které nahrazuje původní RFC 821.

V systému elektronické pošty založené na SMTP protokolu se vyskytují dva logické programové prvky:

- MTA (Message Transport Agent): program určený k předávání zpráv dalšímu uzlu na cestě a k podržení zprávy ve „frontě“ v situaci, kdy cesta dále není z nějakých důvodů dostupná. Pod pojmem MTA rozumíme SMTP server, reprezentovaný například programy sendmail, postfix, qmail případně i MS Exchange či Lotus Notes.
- MUA (Message User Agent): program realizující rozhraní mezi uživatelem zprávu komponujícím a odesílajícím. Takový program je obvykle kombinován i s funkcí čtení zpráv, nicméně pro přístup k doručeným zprávám se používají jiné protokoly či metody než SMTP, kteréžto nejsou pro potřeby této statě podstatné.

Typická cesta zprávy začíná v MUA, kde je komponována, následně je zpráva předána protokolem SMTP definovanému nejbližšímu MTA. Ten zjistí, kudy má být doručována pošta pro konkrétního adresáta a předá opět pomocí protokolu SMTP zprávu následujícímu MTA v pořadí. Řetězec MTA zprávu si předávajících může být jednoduchý, ale může mít i více členů. Cílový MTA, takový, který realizuje přímo poštovní schránku adresáta, zprávu přijme a uloží ji do složky došlé pošty adresáta. Z ní si ji adresát nějakým vhodným způsobem vyzvedne a naloží s ní dle vlastního uvážení.

Protokol SMPT je jednoduchý textový aplikační protokol pracující nad transportním protokolem TCP a používající standardní port 25. Definuje několik základních příkazů a definuje několik návratových kódů (odpovědí). Obecně lze říci, že seznam elementárních SMTP příkazů, se kterými si lze vystačit pro odeslání zprávy, je následující:

- HELO: představení se MTA
- Mail from: petr@odnekud.tld: adresa odesílatele zprávy
- Rcpt to: pavel@nekde.tld: adresa příjemce zprávy
- Data: za příkazem následuje vlastní obsah zprávy (včetně hlaviček emailu zobrazovaných při čtení zprávy), ukončený tečkou (.) na samostatném řádku
- Quit

Typická komunikace pomocí protokolu SMTP mezi MUA a MTA, případně mezi dvěma MTA navzájem, může probíhat následovně (řádky označené K značí odesílající stranu – klienta, řádky označené S značí přijímající stranu – server):

```
S: 220 mail.nekde.cz ESMTP Postfix
K: HELO odnekud.cz
S: 250 Hello odnekud.cz
K: MAIL FROM:<petr@odnekud.cz>
S: 250 Ok
K: RCPT TO:<pavel@nekde.cz>
S: 250 Ok
K: DATA
S: 354 End data with <CR><LF>.<CR><LF>
K: Subject: Pokusna zprava
K: From: petr@odnekud.cz
K: To: pavel@nekde.cz
K:
K: Ahoj
K: Posilam testovaci zpravu.
```

```
K: Mej se pekne.
K: .
S: 250 Ok: queued as 12345
K: QUIT
S: 221 Bye
```

V dnešní době je již za standardní považována podpora protokolu ESMTP, rozšiřující škálu příkazů zejména o lepší možnost řídit předávání zpráv a možnost používat zabezpečení a autentizaci. Komunikace se v nejjednodušším případě příliš neliší od výše uvedené, pouze úvodní představení může vypadat například následovně (použitá konvence jako v předchozím případě):

```
S: 220-mail.nekde.cz ESMTP {postfix version and date}
S: 220 NO UCE. {etc., terms of service}
K: EHLO odnekud.cz
S: 250-mail.nekde.cz Helo odnekud.cz [127.0.0.1]
S: 250-SIZE 14680064
S: 250-PIPELINING
S: 250 HELP
```

Klient se představuje příkazem EHLO (přesmyčka HELO), čímž sděluje podporu ESMTP protokolu, a očekává od serveru potvrzení podpory ESMTP protokolu též. V případě, že server sdělí chybu, klient se pokusí použít příkaz HELO a degradovat komunikaci na protokol SMTP. V uvedeném případě server potvrzuje podporu ESMTP protokolu a sděluje, že omezuje maximální velikost přijímané zprávy na 14680064 B.

Návratové kódy lze rozdělit do tří základních skupin. Návratový kód vždy začíná třímístným číslem, určujícím skupinu:

- 2xx OK : kód potvrzující úspěšné přijetí příkazu či dat nebo akceptaci spojení či zaslaného parametru
- 3xx Data: potvrzení přijetí dat zprávy a jejich zařazení do fronty k vyřízení
- 4xx Warning…: dočasná neboli měkká chyba. Po takovém návratovém kódu je komunikace ukončena, ale zdrojový MTA (či MUA) má pokus opakovat později

- 5xx Error: tvrdá chyba, příkaz nelze vykonat, zpráva má být vrácena odesílateli s příslušnou chybou (například cílová adresa neexistuje).

Znalost těchto několika základních příkazů postačuje pro odeslání emailu „ručně“, například pomocí nástroje TELNET, a jeho úspěšné doručení adresátovi. Jak je z popisu patrné, mechanismus neobsahuje žádné zabezpečení a v původním principu nepoužívá ani žádné kontroly správnosti a oprávněnosti údajů. Je tedy snadno možné do zprávy zadat zcela fiktivní adresu odesílatele nebo použít adresu někoho jiného a vydávat se za něj.

Postupem doby došlo k doplnění možnosti autentizace klienta k SMTP serveru pomocí jména a hesla a i šifrování přenosu pomocí protokolu SSL, tyto techniky jsou však použitelné pouze pro ověření uživatelů provozovateli SMTP serveru známých, takových, kteří mají ve stejné síti nějaký účet. K tomuto účelu slouží SMTP-AUTH [13]. Zatímco v ranných akademických dobách Internetu patřilo v podstatě k dobrým mravům povolení i cizího provozu přes vlastní MTA, dnes se to považuje spíš za neodpustitelný zločin. Většina MTA rozlišuje dle adres Mail from: a Rcpt to:, zda zprávu přijmout či odmítnout, a obvyklé chování má přibližně následující schémata:

K doručení přijmou pouze zprávu, která:

- pochází „zevnitř vlastní sítě“
- je určená uživateli v doméně, kterou daný MTA obsluhuje

V některých případech jsou pravidla rozšířena o následující:

- příchozí pošta z vnitřní sítě je předávána pouze spojením navázaným pomocí SMTP-AUTH
- pošta z vnější sítě od vnitřního uživatele je předávána pouze pomocí SMTP-AUTH
- počítače z vnitřní sítě nesmí přímo komunikovat s externími MTA

Ani tato pravidla však na ochranu proti spamu nepostačují.

## 1.2. K čemu vlastně slouží a jak vzniká spam

Spam je vlastně reklamní sdělení. Jde o zprávu, která se koncového adresáta snaží přimět k nějaké akci, ať se již jedná o nákup nějakého konkrétního zboží, či o nějakou činnost (například protivládní demonstraci). To, čím se spam vymyká ostatním způsobům reklamních sdělení, spočívá v jeho „láci“ a možné masovosti. Jinými slovy, pomocí

elektronické pošty (a jak si dále ukážeme, tak i některých jiných elektronických kanálů) lze velmi levně oslovit velké množství cílových osob. Navíc má emailový spam tu příjemnou vlastnost, že na svého adresáta počká. Zadavatel nemusí tedy platit výrobu a distribuci letáků, které si přečte jen někdo, nemusí platit telefonní poplatky a operátory telemarketingu či výrobu a vysílání televizních nebo rozhlasových spotů, které navíc dopadnou pouze na jedince, kteří v konkrétní čas vysílání dané medium sledují.

Spam v elektronické poště se stal skutečným problémem v okamžiku, kdy se jej coby nástroje propagace a marketingu chopily skupiny či jedinci provozující svůj „obchod" řekněme na hranici pravidel a začaly mezi sebou soupeřit silově ve smyslu „čím více a častěji někoho oslovím, tím více vydělám". Dnešní spamování pracuje s účinností v řádu desetin procent oslovených jedinců. Obtěžujícím se spam tedy stal v době, kdy adresátovi do schránky dojde denně větší množství zpráv, o které zjevně nestojí, a kdy jejich třídění a odstraňování musí věnovat nezanedbatelný čas. Masivní spam mimo to může způsobit problém i technického rázu, například zahlcení přenosové linky (z pohledu pronajímatele balastními) daty či přetížení nebo zahlcení MTA.

Mírně specifickou, ale o to nebezpečnější formou spamu jsou zprávy, které se cíleně snaží vypadat důvěryhodně a vydávat se za zprávy od někoho jiného, nejlépe důvěryhodného. Takovou formou spamu může být tzv. phishing, tedy podvržený email, snažící se o vylákání nějakých citlivých a zajímavých informací od koncového uživatele – obvykle se jedná o čísla a PIN kreditních karet nebo přihlašovací údaje k ne příliš silně zabezpečeným elektronickým bankovnictvím nebo platebním systémům. Může se jednat i o mírně specifickou formu viru, který se ke svému šíření nesnaží využít známou zranitelnost emailového klienta, ale důvěryhodným vzhledem a obsahem se snaží přimět přímo uživatele, aby spustil obsaženou přílohu. Mnohokrát se vyskytl email maskující se za bezpečnostní opravu zaslanou společností Microsoft, obvykle obsahující trojského koně, pomocí nějž je následně počítač vzdáleně ovládán a stává se součástí farmy botů používaných k rozesílání dalšího spamu nebo například součástí farmy použité pro DDOS útok.

Spam, proti kterému se chceme chránit, má tedy několik charakteristických vlastností:

- konkrétní zpráva se na internetu vyskytuje ve velkém množství kopií, často stejnou zprávu dostává jeden uživatel opakovaně

- snahou odesílatelů je rozeslat s minimálními náklady zprávu na maximální počet koncových adres. Tato vlastnost je zvlášť důležitá, neboť se k danému cíli často využívají tzv. zombie počítače – ovládnuté cizí počítače připojené pevně k Internetu. Aby ovládnutí počítače a vykonávaná činnost nebyla odhalena, přitom aby taková činnost probíhala co nejefektivněji specializované odesílací programy neimplementují vůbec a nebo implementují nekorektně určité mechanismy doručování zpráv.

# 2. ZPŮSOBY OCHRANY PROTI EMAILOVÉMU SPAMU

Proti spamu se lze bránit několika skupinami technik dělitelnými dle toho, kdo a v jaké fázi komunikace ochranu realizuje. Jde o:

- Ochranné techniky koncových uživatelů, zaměřené zejména na rozumné chování a přiměřenou úroveň ochrany důležitého údaje, tedy emailové adresy
- Ochranné automatizované techniky na straně infrastruktury emaily přijímající, tyto lze dále rozdělit ještě na dvě podkategorie:
    - odhalit spamera v okamžiku, kdy dochází k navázání SMTP spojení na cílový MTA a nepřijmout zprávu vůbec (zkoumat tedy korektní chování implementace přenosu)
    - odhalit spamovou zprávu po přijetí analýzou obsahu a naložit s ní automaticky dle zadaného pravidla. (snažit se tedy nahradit ruční práci uživatele).
- Ochranné techniky na straně infrastruktury emaily odesílající, zejména snahu zabránit zneužití vlastní techniky pro odesílání spamu a ztížení zneužívání vlastních emailových adres jako podvržených odesílatelů spamu.

Každá z uvedených skupin technik má své předpoklady, vlastnosti a nějakou více či méně měřitelnou úspěšnost. U každé lze nalézt řadu výhod i nevýhod a v praxi je poměrně běžné, že cesta k rozumnému výsledku vede skrz vhodnou kombinaci více technik. Jednotlivé dnes známé a používané techniky budou podrobněji diskutovány v následujících kapitolách.

## 2.1. Ochranné techniky koncových uživatelů

Skupina ochranných technik koncových uživatelů je nejvíce založena na chování koncových uživatelů, dodržování určitých pravidel práce s elektronickou poštou a nakládání s emailovou adresou svojí, případně s emailovými adresami dalšími (firemními, adresami komunikujících partnerů). Ve zbytku kapitoly hovoříme o technikách pracujících s doručeným spamem.

Pro úspěšné rozesílání spamu je třeba mít k dispozici potřebnou infrastrukturu (obvykle farmy počítačů napadených škodlivým kódem a ovladatelných na dálku) a pak zejména emailové adresy potenciálních příjemců. Adresy jsou obvykle získávány dvěma cestami:

- Dolováním (tzv. harvesting) existujících emailových adres z dostupných zdrojů. Jde zejména o WWW stránky, na kterých se vyskytují kontakty, archivy emailových konferencí a podobně. Dále lze adresy získávat pomocí škodlivého kódu, například počítačového viru, který sám sebe rozesílá na veškeré adresy nalezené na napadeném počítači a jako přídavek může tyto adresy předat do spamerské databáze.
- Náhodným generováním adres, obvykle založeným na slovníkovém principu. Spammer jednoduše vygeneruje velké množství adres, přesněji řečeno částí před @, a pokusí se na takto vytvořené adresy v nějaké vhodné doméně odeslat emaily. Pravděpodobně bude většina odmítnuta, ale část adres bude existovat a zprávy budou doručeny.

### 2.1.1. Utajování emailové adresy

Jednou z často používaných technik je utajování reálné emailové adresy tak, aby nemohla být snadno nalezena vyhledávacími roboty a zařazena do spam databáze. Utajování má obvykle různé formy, lze používat technické utajování a časté měnění emailových adres.

Technicky lze emailové adresy utajovat zejména tak, že www stránky nebudou obsahovat přímo odkazy typu mailto:referent@firma.cz, ale tyto budou nahrazeny buď různými zastíracími náhradami (jako jsou například obrázky s uvedenou emailovou adresou) či kódovány do podoby HTML entit, nebo budou doplňovány na základě skriptovacího prostředku na straně klienta WWW. Takové techniky v naprosté většině případů ztěžují použitelnost pro uživatele WWW stránek, emailové adresy není možné jednoduše kopírovat, nelze na ně přímo kliknout. Uvedené techniky tedy jdou proti zásadám přístupnosti a použitelnosti webu.

Další, o něco korektnější formou utajování emailových adres je nahrazení adres přímo kontaktními formuláři. Zájemce o kontakt tedy neodesílá email, ale vyplňuje formulář, který je do emailové schránky adresáta doručen přímo ze serverové aplikace provozované

na www serveru s prezentací. V tomto případě je zase nesnadné zajistit archivaci zprávy na straně odesílatele, ten by možná měl ve své poště rád kopii odeslané zprávy s přesným datem, časem odeslání a podobně. Vzhledem k tomu, že následná komunikace již obvykle přejde do roviny přímé výměny emailových zpráv, kde už adresy skrýt nelze, není takové chování obvykle příliš efektivní. Navíc je v případě kontaktních formulářů poměrně akutní riziko napadení komentářovým spamem a ochranné techniky proti němu jsou pro koncového uživatele obvykle obtěžující.

### 2.1.2. Časté změny emailové adresy, používání dočasných adres

Někteří uživatelé přistupují k chování, kdy zejména v případě privátních adres provozovaných často na některé freemailové službě adresu opouštějí poté, co na ni začne přicházet neúměrně velké množství spamu, a dále ji již nepoužívají. Toto chování je na první pohled velmi nepohodlné a vede k neustálému přetrhávání navázaných komunikačních partnerství, neboť je stále nutné oznamovat novou adresu svým obvyklým partnerům. Lidé závislí na emailové komunikaci budou ale nejspíš souhlasit s tvrzením, že: neustálé měnění emailové adresy si může dovolit pouze člověk paranoidní, nebo takový, se kterým ve skutečnosti nikdo komunikovat nechce.

Přijatelnější alternativou je použití dočasné adresy v případě potřeby zadat platnou emailovou adresu někde, kde je předpoklad skutečně jednorázového použití a kde uživatel z různých důvodů nechce udat svoji stabilní adresu (obvykle jde o různé registrační formuláře do služeb ve spektru od zřejmě seriózních až po na první pohled pochybné). Dočasné adresy lze získat na mnoha místech, lze je vytvářet na freemailových službách nebo využít například služby mailinator.com, která přijme email na jakoukoliv adresu v domeně mailinator.com a 24 hodin ji udržuje ve schránce, přičemž prohlédnout si ji prostřednictvím www rozhraní může kdokoliv, kdo zná řetězec před zavináčem v adrese.

Další, v některých případech doporučitelnou technikou je využití systémové variabilnosti adres v SMTP protokolu definované RFC, která spočívá v možnosti za označení mailboxu v adrese (do části před znakem @) vložit za speciální znak rozlišovací řetězec. Ten se při doručování do konkrétního mailboxu následně ignoruje. Jako speciální oddělující znak je možné použít znak +, v mnoha případech i -. Adresa petr@firma.cz a petr+registrace1@firma.cz jsou ekvivalentní a takto adresované zprávy budou doručeny do stejného mailboxu. V případě potřeby je ale možné dle celé emailové adresy třídit zprávy

automatickým filtrem, a začne-li tedy na adresu petr+registrace1@firma.cz docházet spam, lze jej snadno třídit a mazat. Jako vedlejší efekt je známo také to, kam byla tato adresa prvotně zadávána, a kdo ji tedy „prodal“ pro spamování.

### 2.1.3. Oznamování spamu

V případě, že koncový uživatel nebo koncová síť je zahlcována příliš velkým množstvím zpráv pocházejících z jednoho zdroje (z jedné IP adresy), je možné postupovat metodou hlášení spamu. Obvyklou první instancí je administrátor místního emailového serveru, který by měl být schopen z hlavičky emailu zjistit skutečného původce zprávy (přesněji řečeno skutečnou IP adresu počítače, ze kterého byla zpráva prvotně odeslána) a může přímo na serveru zablokovat příjem z této adresy. Dále mmůže kontaktovat přímo správce sítě, v níž se zdrojový počítač nachází, se žádostí o nápravu situace nebo blokování odesílání pošty z této adresy, případně může incident nahlásit na některou službu sbírající takové informace. Ta následně zařadí příslušnou IP adresu na blacklist a kontaktuje správce cílové sítě s upozorněním na takové opatření a žádostí o nápravu a reportování řešení problému. Blacklist jako ochrana proti spamu je podrobněji diskutován v kapitole 2.2.9.

### 2.1.4. Zabezpečení počítače koncového uživatele

Mezi metody ochrany proti spamu na straně koncového příjemce elektronické pošty lze zařadit i vhodné zabezpečení počítače uživatele. Pokud bude počítač vhodně nakonfigurován tak, aby nebylo možné nebo snadné využít zranitelnosti emailového klienta (používaní MUA s malým počtem objevených chyb, pravidelná aktualizace a aplikace bezpečnostních záplat…), bude nastaven tak, aby bránil šíření počítačových virů a bude chráněn vhodným a aktualizovaným antivirem, je pravděpodobné, že z něj nebudou získány emailové adresy vhodné pro zařazení do spamerské databáze a že sám nebude použit pro skryté rozesílání spamu.

### 2.1.5. Shrnutí technik koncových uživatelů

Z uvedených kapitol je zřejmé, že techniky použitelné koncovými uživateli jsou velmi omezené, náročné na dodržení pravidel chování a jejich efektivita celkově není příliš vysoká. Jednoznačně lze doporučit přiměřené zabezpečení pracovní stanice koncového uživatele (aktuální antivir, aplikace bezpečnostních oprav, další opatření například v závislosti na firemní bezpečnostní politice…) a dodržování přiměřených opatření při

náhodné nebo ne příliš důvěru budící emailové komunikaci. Jedná se zejména o používání dočasných nebo variabilních emailových adres v případě potřeby skutečně jednorázové komunikace a dále o naprostou zdrženlivost v odpovídání nebo jakýchkoliv reakcích na doručený spam. Na spamové emaily obvykle nelze přímo odpovědět, neboť adresy jsou podvržené nebo neexistují. Reakce prostřednictvím odkazů nebo adres uvedených přímo v těle zprávy je často považována za potvrzení „života“ cílové schránky a vede s větší pravděpodobností k zvýšení intenzity doručovaného spamu než k jeho snížení.

Jednoznačně nejefektivnější v tomto směru je nasazení vybrané kombinace automatizovaných technik ochrany proti spamu, které budou diskutovány v následující kapitole.

Není vhodné ani doporučeníhodné důsledné utajování emailových adres. Většina těchto technik pouze komplikuje život potenciálním partnerům, o které máme ve skutečnosti zájem, a s mnoha z nich si dolovací roboti dokážou poradit.

## 2.2. Automatizované ochranné techniky z pohledu infrastruktury emaily přijímající

### 2.2.1. Pravidlová analýza obsahu

Začněme technikou založenou na analýze přijaté zprávy nějakým automatem, který podrobí doručenou zprávu testu dle mnoha dostupných pravidel a každé pravidlo hodnotí pomocí definovaného „skóre“. Ve výsledku se do celkového skóre započtou veškeré dílčí výsledky. Tímto způsobem funguje například oblíbený program spamassasin [2], často nasazovaný na MTA provozovaných na systémech Unix.

Metoda pravidlové analýzy je založena na tom, že zkoumá mnoho charakteristických pravidel a při jejich splnění připočte k definovanému skóre příslušnou váhu. Po překročení nastavené hodnoty je zpráva vyhodnocena jako spam a je takto označena. Pravidla jsou vytvářena komunitou podílející se na vývoji programu a jsou neustále doplňována a aktualizována. Do programu je pochopitelně možné doplňovat vlastní pravidla zohledňující specifické podmínky sítě, v níž program nasazujeme, nebo například specifické podmínky jazyka obvykle používaného pro komunikaci v místě nasazení a odlišného od angličtiny.

Pravidla lze definovat a nasazovat na vyhodnocení těla emailu (výskyt konkrétního slova, neobvyklé formátovaní nebo neobvyklý obsah těla emailu) a lze je též definovat pro hlavičky emailu, a to jak standardní, tak speciální. V obou případech lze využívat regulárních výrazů a získat tedy velmi flexibilní nástroj k definici pravidel.

Mimo jednoduchá pravidla uplatňovaná na hlavičky a tělo vlastního emailu existuje ještě možnost definovat speciální typy pravidel, a to meta, URI a rawbody. Jejich rozšířená funkcionalita spočívá v případě URI pravidel ve vyhodnocování výhradně URL obsažených v případné HTML části zprávy, čímž je umožněno efektivněji vyhledávat odkazy na spamerské www prezentace a hodnotit je významnějším skóre. Rawbody pravidlo umožňuje vyhodnocovat tělo emailu přesně ve stavu, v jakém bylo převzato, neprovádí se tedy odstranění například HTML značek nebo interpretace MIME sekvencí. Meta pravidla jsou určena k vyhodnocování kombinací jiných pravidel, lze tedy určitá pravidla spojovat do logických nebo aritmetických celků a jejich výsledné skóre přidělovat na základě splnění celého celku. Vzhledem k tomu, že skóre může být i záporné, lze meta pravidla využívat například k zachycení a správnému vyhodnocení zpráv, které sice splňují mnoho obvyklých charakteristik spamu, ale v našem konkrétním případě jde o korektní zprávu, která musí být doručena. Mezi jednoduchá pravidla může patřit například splnění následujících podmínek (s odůvodněním):

- adresa odesilatele je z domén yahoo.com či msn.com (tyto domény byly často využívány pro fiktivní adresy odesílatelů spamu, jako svého času největší freemailové služby nevyvolávaly podezření)
- jméno odesílatele obsahuje číslo (mezi spamery často používaná praktika).
- zpráva obsahuje pouze přiložený obrázek
- interní formát zprávy není korektní (případně obsahuje nějakou konkrétní známou chybu).
- IP adresa, odkud zpráva přišla, nepatří do sítě uváděného odesílatele
- adresa odesílatele je z vlastní vnitřní sítě, ale zpráva přišla „zvenčí“
- …

Pravidlové systémy používají i složitější pravidla, která například některé charakteristiky vyhodnocují v aktuální okamžik proti zdrojům dostupným v síti nebo podrobují složitějšímu zkoumání. Součástí pravidel tedy může být podrobení obsahu emailu lexikální

analýze pomocí Bayesova teorému nebo například vyhledání odesílající IP adresy v sadě na Internetu udržovaných blacklistů (obě techniky jsou samostatně vysvětleny dále v textu).

Jak je patrné, pravidel může být mnoho a jejich stanovení záleží zejména na empirické zkušenosti administrátora nebo provozovatele systému. V praxi je situace obvykle taková, že součástí programu na pravidlovou analýzu obsahu je sada možných pravidel a ten, kdo systém nasazuje, sám přiřadí pravidlům příslušné váhy. Mezi výhody takového systému patří relativní jednoduchost instalace a provozování, mezi nevýhody naopak patří:

- nutnost přenést celou zprávu až k adresátovi (čerpat tedy přenosové pásmo linky)
- nutnost zprávu analyzovat (čerpat tedy strojový čas)
- obecná známost používaných pravidel, spameři se tedy mohou pravidlům přizpůsobovat. Systém je třeba neustále upravovat, jinak ztrácí účinnost.
- hrozba „false positives“, tedy korektních emailů označených jako spam

### 2.2.2. Lexikální analýza obsahu zprávy (Bayes)

Metody lexikální analýzy, obvykle využívající Bayesův teorém [3], spočívají v porovnání četnosti výskytů určitých řetězců ve zprávách označených jako spam a také ve zprávách označených jako ham, tedy korektních zprávách. Tyto metody tedy vyžadují „učení“, je nutné předložit jim před použitím dostatečně velký vzorek obou typů zpráv. I takovéto metodě se lze bránit, obvykle tak, že jsou zaměňovány znaky v problematických slovech tak, že pro počítač jde o jiné, nezávadné řetězce, ale živý čtenář je správně přečte (v1agra, v i a g r a, …). Tato metoda bývá v současné době zabudována v oblíbených programech pro práci s elektronickou poštou (Mozilla Thunderbird, …) nebo dostupná prostřednictvím samostatných programů (iFile). Bývá též součástí pravidlového systému, tedy přispívá svým výsledkem do celkového skóre. Výhodami systému založeném na Bayesovském filtrování je jednoduchost jeho zprovoznění, není třeba nastavovat velké množství pravidel. Mezi nevýhody lze zařadit:

- nutnost systém naučit a průběžně doučovat oba typy zpráv.
- nutnost přenést celou zprávu až k adresátovi (čerpat tedy přenosové pásmo linky)
- nutnost zprávu analyzovat (čerpat strojový čas)
- obecná známost metody, snaha spamerů měnit problematická slova a systém tak obcházet
- hrozba „false positives“, tedy korektních emailů označených jako spam.

### 2.2.3. Ověření kontrolního součtu

Princip této metody spočívá v předpokladu, že většina spamu se vyskytuje na internetu ve velkém množství identických kopií, lišících se pouze hlavičkami (adresa příjemce, podvržená adresa odesílatele) a případně drobnostmi v obsahu emailu, jako například oslovením. V takovém případě lze z vlastního obsahu zprávy odebrat potenciálně rozdílné texty a ze zbytku spočítat vhodnou hash funkcí kontrolní součet. Ten lze porovnat s databází kontrolních součtů zpráv označených již dříve jako spam a v případě shody zprávu přímo vyřadit nebo zvýšit její spam skóre. Pokud v databázi není a uživatel přesto zprávu vyhodnotí jako spam, lze kontrolní součty na rozdíl od ostatních informací do databáze snadno a rychle reportovat. Tím získáváme stav, kdy reporting případných zdrojů spamu není závislý na správcích, musejících odhalit IPadresu původce spamu a správným způsobem ji reportovat do databází blacklistů, ale spam může oznámit snadno jakýkoliv uživatel, aniž by byl vybaven jakýmikoliv technickými znalostmi. Komunita spolupracující na identifikaci spamu se tak snadno rozšíří o několik řádů a v případě nově se vyskytnuvšího masově šířeného sdělení je šance, že jej někdo z postižených velmi brzy nahlásí a ostatním od něj tak odpomůže.

### 2.2.4. Nevýhody systémů založených na analýze obsahu

Jak je vidět, systémy založené na analýze obsahu zprávy trpí společnými nevýhodami, mezi které patří zejména:

- Nutnost přijmout celý email. Při masivním útoku na server s větším množstvím uživatelů to může vést k zahlcení linky, případně k čerpání placeného datového toku, následně spamy spotřebovávají místo na disku.
- Nutnost zprávu analyzovat a věnovat tomu tedy strojový čas. Při masivním útoku může náročnost zpracování zpráv přetížit server tak, že znemožní či velmi zpomalí provoz korektní pošty.
- Hrozba „false positives“, tedy nesprávně vyhodnoceného typu zprávy.

### 2.2.5. Metody odhalení spamu před přijetím dat

Metody odhalení spamu před přenesením dat se snaží pomocí využití údajů ze základní SMTP komunikace a určitého, někdy i záměrně vyvolaného, chování odesílajícího serveru odhalit nekorektně se chovající odesilatele a v případě, že je rozeznají jako zdroj spamu, od

nich data vůbec nepřijmout. Využívá se obvykle toho, že se spameři snaží odeslat co nejvíce zpráv za co nejmenší cenu a neimplementují korektně veškerá pravidla SMTP přenosů.

Čeho lze vlastně využít k rozeznání spamera:

- IP adresy odesílajícího serveru
- lze zjistit, zda IP adresa odesílatele má korektní PTR záznam v DNS a ten se zpětně resolvuje korektně
- adresy Mail from:
- adresy Rcpt to:
- dalších záznamů přístupných přes externí zdroje (DNS, externí databáze)
- technického chování odesílajícího serveru

### 2.2.6. Test existence a korektnosti PTR záznamu v DNS

V tomto případě se obracíme na DNS systém, který obsahuje překlady IP adres na doménová jména a opačně. Dobré mravy a částečně technická nutnost předepisují, aby každé definované doménové jméno bylo přeložitelné na IP adresu pomocí záznamu A nebo CNAME a každá IP adresa měla přidělené doménové jméno pomocí záznamu PTR. Při zahájení SMTP komunikace známe IP adresou odesilatele. Lze provést pokus přeložit ji pomocí DNS na doménové jméno (ověřit PTR záznam) a tento při jeho existenci opět převést na IP adresu. Pokud se podaří kruh dokončit, měli bychom získat stejnou IP adresu jako na začátku komunikace. Pokud dostaneme jinou IP adresu nebo PTR záznam neexistuje vůbec, lze odesílající MTA považovat za podezřelý. Bohužel i u poměrně velkých ISP dochází k situaci, kdy uvedený test neprojde a přesto nelze poštu odmítnout. Výhodou tohoto testu je jeho jednoduchost, nevýhodou nespolehlivost. Lze ho použít jako dodatečné pravidlo pro pravidlové systémy.

### 2.2.7. Ověření zpětné adresovatelnosti odesílatele

Tato metoda spočívá v tom, že po přijetí Mail from: hlavičky se přijímající MTA připojí na MX server domény z Mail from: a zadá obdrženou adresu jako Rcpt to:. Pokud neobdrží kladné vyjádření (250 OK), odmítne přijmout zprávu, neboť adresu odesílatele považuje za podvrženou. V opačném případě zprávu přijme, neboť protistrana je pro odesílatele ochotná zprávu též přijmout.

Metoda ověření zpětné adresovatelnosti odesílatele bohužel velmi omezeně použitelná, nevýhody spočívají zejména ve zdvojnásobení navazovaných TCP spojení a tím zátěže pro přijímající MTA, dále v tom, že je bohužel běžná praxe odesílat zprávy z neadresovatelných adres. Navíc odmítnutí může nastat například i z důvodů použití jiné antispamové ochrany na testovaném protiserveru, například Greylistingu (vysvětleno v kapitole 2.2.11). Mezi další omezení lze obecně započíst již skutečnost, že se spoléháme na prostředek nebo zdroj, který nemáme pod kontrolou, může být nedostupný nebo vytížený, a proto bude dlouhými odezvami ovlivňovat provoz našeho vlastního systému. Dalším rizikem je to, že ověření zpětné adresovatelnosti odesilatele projde v případě, že testovaná doména používá doménový koš (přijme email na jakoukoliv adresu), což též není úplně neobvyklé chování.

### 2.2.8. Test korektní implementace SMTP

Tato metoda využívá v RFC definovaných pravidel a požadavků na úvodní komunikaci, zejména těch, které se dotýkají příkazu HELO/EHLO. První vlastnost, se kterou lze provést test, je požadavek RFC, aby odesílající strana zaslala příkaz HELO/EHLO až poté, co obdrží od přijímající strany úvodní hlášení (banner). Odeslání banneru lze pozdržet a testovat, zda počká i strana odesílající. Spamovací jednoúčelové stroje obvykle ve snaze urychlit rozesílání spamu na banner kód nečekají a zahájí komunikaci bezprostředně po navázáni TCP spojení. Nevýhodou tohoto chování je zdržování komunikace netriviálním způsobem, metodu lze obecně využívat pouze pro testování podezřelých odesílatelů plynoucí z jiných testů.

Ve stejném okamžiku se stejným účelem můžeme testovat, zda je příkaz HELO zasílán korektně, se správným formátováním, zda je v příkazu HELO uváděno existující FQDN a zda odpovídá IP adrese, ze které je navázáno spojení. Spam servery často v HELO uvádějí zcela neplatné nebo podvodné údaje, které je ale možné snadno ověřit.

### 2.2.9. Black/white listy

Principem blacklistů a whitelistů je vyhledávání IP adresy odesílajícího MTA na nějakém vhodně zvoleném nebo vlastními silami budovaném seznamu povolených nebo naopak zakázaných IP adres. S požadavkem na spojení je naložena dle výsledků tohoto porovnání.

Blacklist představuje jednu z prvních metod obrany proti spamu používanou na Internetu. Mechanismus je založen na tom, že komunita buduje seznamy MTA, ze kterých byl odeslán spam, případně takových, kde je MTA nevhodně nakonfigurován (povolený OpenRelay). Problém spočívá zejména v tom, že na určitý blacklist se může náš MTA dostat prostě proto, že někdo z našeho serveru obdržel email, který považoval za spam, a zapsal náš server na blacklist. Následně od nás nekontrolovatelné množství MTA přestane zcela odebírat zprávy, a to bez ohledu na to, zda zápis byl vůbec proveden na základě relevantního provinění, případně na to, jaký poměr spamu a korektních zpráv náš systém v čase odesílá. Použité informace čerpají z článků Nepoužívejte IP blacklisty [4] a [5].

Blacklisty technicky fungují totožným systémem jako DNS, jsou často označovány zkratkou DNSBL. Bývají provozovány pomocí stejného software jako běžné DNS servery a pro ověřování záznamú odesílajícího MTA na blacklistu jsou používány způsoby reverzních dotazů do DNS. Prakticky se tedy postupuje následujícím způsobem:

- vezmeme IP adresu odesílajícího MTA (například 192.168.1.13) a otočíme ji (tedy 13.1.168.192)
- připojíme doménové jméno DNSBL seznamu, ve kterém chceme adresu ověřit (například 13.1.168.192.spam.seznam.com)
- provedeme do DNS systému dotaz na existenci takového záznamu (vyhledáváme tedy záznam typu A). Obdržíme buď IP adresu a tedy informaci, že uvedená IP adresa na seznamu je, nebo odpověď NXDOMAIN, tedy neznámá doména.
- v případě pozitivního nálezu můžeme zkusit ověřit TXT záznam, který obvykle obsahuje informaci, proč byla IP adresa na seznam zařazena.

Popis fungování DNSBL lze nalézt v Encyklopedii Wikipedia [14] a seznam a vlastnosti často používaných DNSBL v Encyklopedii Wikipedia [15].

### 2.2.10. Habeas SenderIndex a SafeList

SenderIndex a SafeList jsou komerční produkty bezpečnostní a konsultační společnosti Habeas [8]. Jde ve své podstatě o sofistikovaný způsob kategorizace a klasifikace IP adres a domén do whitelistů a blacklistů na základě certifikace, případně sledování specialistů společnosti Habeas. V rámci programu SenderIndex je možné projít certifikací společnosti Habeas a získat zařazení na SafeList, volně přístupný whitelist IP adres, které lze

považovat za důvěryhodné - zprávy od takových IP adres přijaté není nutné dále zkoumat a klasifikovat ve snaze o odhalení spamu. Mimo to je dostupný AcceptList, seznam důvěryhodně vystupujících IP adres na Internetu, které však neprošly certifikací společnosti Habeas, a je tedy vhodné od nich zprávy obdržené zkoumat dalšími antispamovými nástroji, a také BlockList, blacklist IP adres, od kterých je vhodné zprávy nepřijímat.

### 2.2.11. Greylisting

Tato metoda je založená na jakémsi dynamicky budovaném vlastním seznamu povolených charakteristických kombinací informací emailových zpráv. Následující popis čerpá z [6] a [7]. U každé zprávy zaznamenáváme trojici údajů: IP adresa odesílajícího MTA, Mail from: a Rcpt to: ze SMTP komunikace. Následně ve vlastní databázi zjišťujeme, zda se v minulosti tato kombinace již vyskytla - v kladném případě email přijmeme, v opačném případě provedeme záznam těchto údajů s časovou značkou do databáze a zprávu odmítneme s chybovým kódem 450. Ten znamená dočasnou chybu a dle RFC má odesílající MTA pokus o doručení zprávy po nějakém čase zopakovat. Korektní odesílající MTA toto udělá, a protože MTA na straně příjemce již má trojici informací uloženou ve své databázi, zprávu přijme. Spameři však obvykle pracují s velmi nespolehlivými seznamy adres, obsahujícími mnoho neplatných adres, případně adresy dokonce náhodně generují, a z toho důvodu obvykle chyby neřeší, a to ani chyby dočasné 4xx. Mimo to se zprávy obvykle rozesílají z ovládnutých počítačů, které ve snaze neprozradit svoji přítomnost nemohou spotřebovávat dost procesorového času na evidenci neplatných pokusů a snažit se o znovudoručení později.

Tato metoda má své odpůrce, charakterizující její nevhodnost zejména tím, že příjemce pošty přesouvá svůj potenciální problém se spamem na cizí zdroje, tedy zdroje odesílatele pošty. Používání greylistingu má k několika nepříjemným důsledků:

- Zpožďování emailů. Při prvním pokusu o komunikaci mezi dvěma emailovými adresami dojde ke zdržení zprávy na předem neodhadnutelnou dobu. Doba záleží na prodlevě, po které odesílající MTA zkusí komunikaci zopakovat. Tato prodleva je dynamická a závisí na nastavení odesílajícího MTA a například i na jeho vytížení. Na druhou stranu je email považován za obdobu klasické papírové pošty a negarantuje dobu doručení zprávy, proto tato vlastnost není v rozporu s RFC 2821 [12] ani principy jeho fungování.

- Možnou změnu pořadí zpráv. V případě, že při prvotním kontaktu mezi dvěma emailovými adresami je v relativně krátkém intervalu postupně odesláno z adresy A na adresu B více emailů, dojde pravděpodobně k situaci, kdy první email v pořadí bude doručen až jako druhý nebo další v pořadí proto, že bude odesílajícímu MTA odmítnut, ale druhý v pořadí čekající ve frontě bude přijat a doručen jako první, bude-li odeslán po uplynutí v greylistingu nastavené prodlevy, neboť vyznačuje stejnou trojicí charakteristických informací. Toto chování může být pro příjemce poněkud matoucí a nepochopitelné.
- Ve specifických situacích může dojít k velmi velkému zpoždění doručení zprávy. Problém nastane v případě, že odesílající strana je síť nebo poskytovatel služeb s velkým SMTP provozem, který pro odesílaní využívá farmu SMTP serverů s různými IP adresami. V některých případech pak odmítnuté zprávy může opakovaně zkoušet doručovat jiný SMTP server (tedy s jinou IP adresou) než při prvním pokusu a mail je opět odmítnut. Prodleva, než zprávu zkusí podruhé odeslat některý ze serverů, který se o to již pokusil, může být poměrně velká. Tímto chováním se vyznačuje například GMail, ale i některé další freemailové služby. Riziko lze odstranit whitelistováním vybraných IP adres známých farem odesílajících emaily, nebo odlehčenou verzí greylistingu, kde zaznamenáváme pouze prvních 24 bitů IP adresy odesílajícího serveru, a farma sdílející stejný C rozsah IP adres je tedy považována za jeden zdroj zpráv.
- Problém nastává i v případě některých diskusních skupin provozovaných prostřednictvím emailů, které k lepšímu sledování doručených a nedoručených zpráv používají unikátní emailové adresy odesílatele. Generují je za pomoci zvláštních znaků (+ nebo – v názvu mailboxu). Pokud toto použitý greylistovací engine neimplementuje, zdrží z takového listu každý příchozí email, což při více uživatelích takových diskusních skupin a větším provozu může snadno generovat velké množství záznamů v databázi a vyčerpávat zbytečně výpočetní kapacitu systému příjemce.
- Jeden z často uváděných argumentů proti greylistingu je ten, že v okamžiku jeho většího rozšíření nastane situace, kdy spameři budou veškeré emaily odesílat dvakrát. K tomu zatím nedošlo, i když greylisting je již využíván poměrně dlouho a není okrajovou záležitostí. Očekávané reakci spamerů se lze navíc celkem

účinně bránit zvýšením úvodního intervalu, po který email není přijat. V současné době postačuje prodleva 60 vteřin a je obvykle nastavována při nasazení greylistingu. Pokud by bylo potřeba, lze tuto prodlevu zvětšit například až na 60 minut, nebo i více, a je velmi nepravděpodobné, že spameři budou případný opakovaný pokus provádět s takovým zpožděním (mj. proto, že odesílající počítače bývají v čase velmi nestabilní).

Přes uvedené nevýhody je v dnešní době greylisting jako metoda boje proti spamu velmi účinný, dokáže odmítnout přenos vysokého procenta spam zpráv a jeho hlavní a velmi velkou výhodou je to, že v podstatě netrpí na false-positives. Pokud odesílající MTA korektně implementuje SMTP dle RFC, nehrozí, že by při použití greylistingu nebyl doručen korektní email. Greylisting proto lze doporučit jako jednu z prvních metod v řetězci opatření proti spamu na straně příjemce elektronické pošty, neboť tím, že dokáže s vysokou účinností a za relativně malých nákladů odfiltrovat spam následně velmi zvyšuje efektivitu dalších opatření, například nasazení pravidlové nebo lexikální analýzy obsahu zpráv. Vzhledem k malému množství ponechaných spamů zlepšuje i účinnost navazujících metod. Protože se i počítačové viry samovolně se šířící pomocí emailu vyznačují stejně nízkým respektem k implementování SMTP dle RFC, nasazení greylistingu dokáže zadržet i značnou část virů doručovaných na adresy chráněné sítě a tím řádově snížit zátěž nasazeného antivirového řešení. Mimo to dokáže zastavit i některé virové útoky vedené viry, které nejsou v daný okamžik známy a které nasazené antivirové řešení neodchytí.

### 2.2.12. SPF a SenderID

Metoda SPF a SenderID je založená na tom, že správce domény může prostřednictvím DNS systému umístit do TXT záznamu domény vhodným způsobem popsaná pravidla, jež určují, jaké počítače mohou z jeho domény odesílat poštu s adresami obsahujícími danou doménu. Na straně přijímajícího MTA lze zkontrolovat existenci pravidel, a pokud existují, vyhodnotit, zda jim odesílaná zpráva vyhovuje. Vzhledem k tomu, že spameři odesílající z vlastní domény si SPF nebo SenderID snadno nastaví, kladný průchod nelze úplně vyhodnotit jako pozitivní údaj o hamu, lze ale odmítat emaily, které nevyhoví. Metoda vlastně není antispamovým opatřením sama o sobě, snaží se pouze znemožnit podvrhování

adres odesílatele, zejména na hodnoty některých dobře známých a spamery zneužívaných názvů firem, jako jsou Microsoft, Yahoo a podobně.

SPF ve verzi 1 definuje osm mechanismů kontroly, každý může být kvalifikován jedním ze čtyř kvalifikátorů.

Mechanismy kontroly jsou následující:

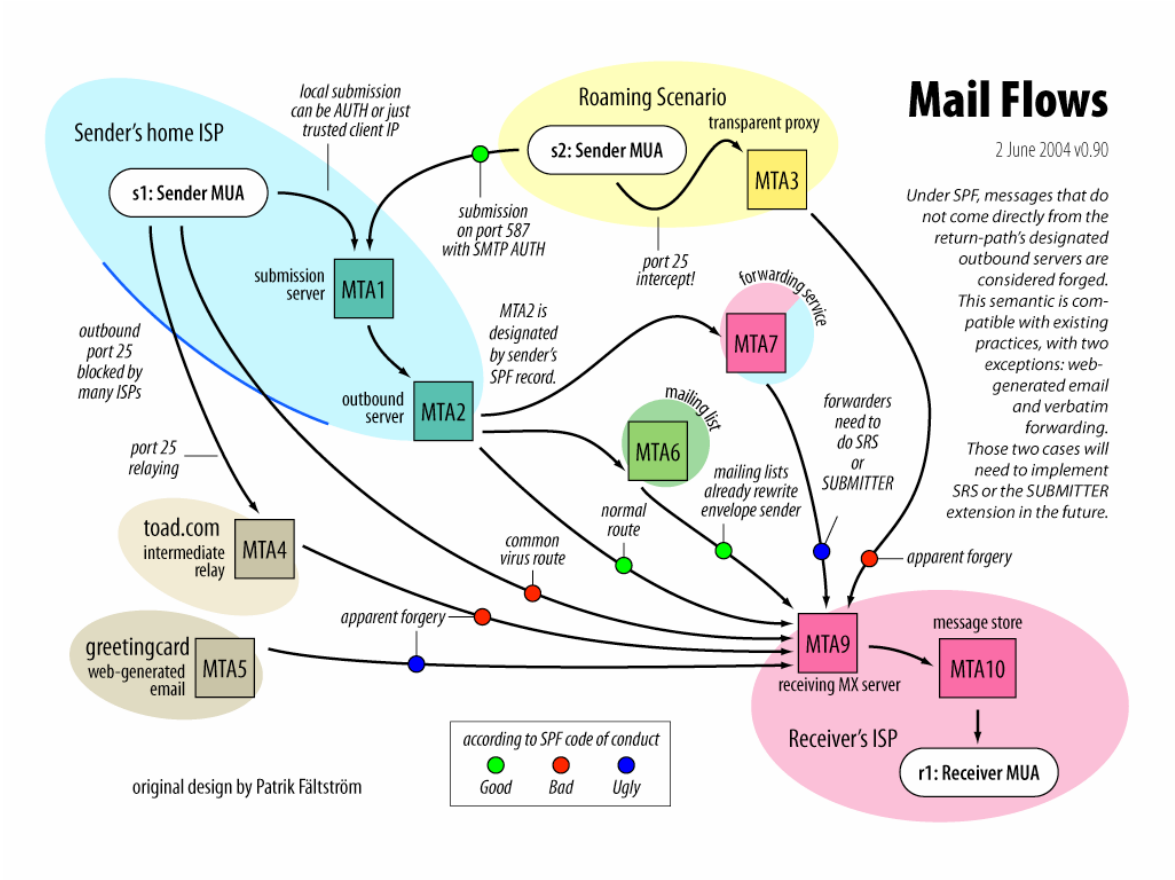
- ALL - toto kritérium je splněno vždy (obvykle se používá jako koncové pravidlo s kvalifikátorem -, který způsobí neplatnost takového pravidla)
- A je platným mechanismem, pokud odesílající doména má platný A záznam v DNS
- IP4, IP6 je platným mechanismem, pokud odesílající IP adresa odpovídá dané IP adrese nebo rozsahu adres protokolů IP (IPv4) nebo IPv6
- MX je platným mechanismem, pokud odesílající IP adresa nebo jméno je definované jako MX záznam v DNS
- PTR je platné, pokud odesílající IP adresa má platný PTR záznam v DNS příslušné odesílající domény
- EXIST umožňuje pouze otestovat existenci záznamu pro doménu v DNS bez ohledu na to, zda odpovídá IP adrese
- INCLUDE je použito pro vložení SPF pravidel použitých pro jinou doménu. Lze tak snadno definovat pravidla pouze jednou i pro větší počet hostovaných domén.

Kvalifikátory jsou následující:

- + je kladný kvalifikátor, jeho použití je implicitní (tedy +MX a MX je totožné). Pokud je pravidlo označené tímto kvalifikátorem splněno, je výsledek kladný
- - je záporný kvalifikátor, při splnění pravidla označeného – je výsledek celého testu záporný a zpráva je odmítnuta
- ? je neutrální kvalifikátor, při splnění takto označeného pravidla není zpráva označena ani jako korektní, ani nekorektní
- ~ měkký záporný kvalifikátor. Při splnění takového pravidla je zpráva přijata, ale pravidlo je nesplněno. Používá se zejména pro ladění nově navržených pravidel.

Nevýhody této metody spočívají zejména v malém počtu domén propagujících SPF nebo SenderID záznam. Mimo to dochází k problémům s korektním přeposíláním emailů (forward), případně využíváním služeb typu WWW přání, kde lze vybrat obrázek s přáním, připojit komentář a emailovou zprávou z adresy zadavatele upozornit příjemce na existenci přání pro něj. V takovém případě je třeba implementovat SRS nebo obdobnou službu, která vhodným způsobem přepisuje hlavičky SMTP komunikace tak, aby odpovídaly reálnému technickému zdroji emailu.

Možné cesty směrování emailů popisuje obrázek 1 převzatý z [17] http://old.openspf.org/mailflows-l.png




Obrázek 1: Možné cesty směrování emailových zpráv

### 2.2.13. Pasti

Jednou z metod boje proti spamu mohou být i jakési pasti. Lze například založit „neplatné“ emailové adresy, tedy takové, které nejsou a nikdy nebyly používány a nebyly nikde propagovány s výjimkou skrytých www stránek, případně byly záměrně zadány do spam

seznamů (například přihlašovací formuláře pochybných WWW služeb). Odesílající MTA, který se pokusí o doručení zpráv na takové adresy, může být následně celkem bezpečně označen za spamera a zprávy z něj mohou být odmítány.

Obdobnou pastí může být falešný MTA, obvykle propagovaný v DNS s velmi nízkou prioritou. Předpokládá se, že dle RFC musí odesílající MTA použít pro doručení pošty MX server s nejvyšší prioritou a další použít pouze v případě, že tento je prokazatelně nedostupný. Lze tedy například nastražit terciální MTA na jednom stroji, ale dvou různých IP adresách jako primární (tak aby byly vždy dostupné oba nebo žádný). Pak lze předpokládat, že pokus o doručení zprávy na takovýto MTA je spam a je možné například velmi zpomalovat přenos dat a tím spamera neúměrně zdržovat. Technika není příliš využívána, neboť příjem spamu na běžné MTA nesnižuje významným způsobem a víceméně plýtvá zdroji jak spamera, tak příjemce.

### 2.2.14. Metody zpoplatňování zpráv

Obecně jde o skupinu technik v současnosti nepoužívanou, pouze v některých případech teoreticky diskutovanou. Princip je podobný jako v běžné papírové poště, tedy zpoplatnit odeslání zprávy buď reálnou hodnotou (tedy penězi) v některých případech modifikovanou tak, že zpoplatnění je provedeno pouze v případě, že koncový uživatel označí zprávu jako spam, nebo zpoplatnit odesílanou zprávu například strojovým časem odesílatele, tedy požadavkem na provedení netriviálního výpočtu, jehož výsledek lze na straně příjemce snadno zkontrolovat. Uvedené metody narážejí na nutnost velkých zásahů do SMTP protokolu, a tedy v podstatě budování paralelního emailového systému. Pravděpodobnost jejich budoucího úspěchu je poměrně malá.

### 2.2.15. Shrnutí automatizovaných technik působících na straně příjemce emailu

V předchozích kapitolách jsme popsali techniky založené na několika principech:

- analýza a hodnocení obsahu zprávy
- vyhodnocení „kreditu“ odesílatele buď na primitivním principu typu přítomnost na zakázaném či naopak privilegovaném seznamu, nebo komplexnější hodnocení kreditu odesílatele komunitou nebo nějakou komerční službou

- jednoduché technologické testy, zjišťující korektní nastavení infrastruktury odesílatele dle příslušných RFC
- testy korektní implementace SMTP protokolu dle příslušných RFC
- techniky založené na principu zpoplatňování zpráv

Z pohledu efektivnosti a nároků na strojový čas a datové kapacity přijímající strany jsou výhodné techniky nevyžadující přenos celé zprávy. Takové testy mohou být v některých případech velmi rychlé a nenáročné, ty nejprimitivnější bohužel nejsou nejspolehlivější. Vzhledem k požadavku na bezpečnost, přiměřeným nárokům na výkon a potřebě malého sklonu k falešnému označování korektních zpráv je obvykle budován řetězec testů, začínající primitivními testy vyřazujícími pouze nejzjevnější spamy a končící kontrolou obsahu zpráv předchozím řetězcem propuštěných.

## 2.3. Ochranné techniky z pohledu infrastruktury emaily odesílající

Každá síť nebo infrastruktura, která pro své uživatele emaily přijímá, je též odesílá. Odesílá emaily ze svých serverů, na které je vlastní uživatel dopravují ze svých MUA spuštěných na počítačích uvnitř vlastní sítě nebo i vně této sítě.

Správce a vlastník infrastruktury by měl v rámci ochrany proti spamu zajistit i ochranu před rizikem, že se jeho vlastní síť stane zdrojem spamu, a vůči této síti a jejím zdrojům následně budou podniknuta opatření externími subjekty (například zahrnutím do některého blacklistu), případně že se vlastní síť stane zdrojem virové nákazy. Mimo to je žádoucí pokusit se nastavit taková pravidla a podmínky, aby ani emailové adresy spadající pod řízenou infrastrukturu nemohly být snadno zneužívány jako podvržené adresy, ze kterých pochází spam.

Z pohledu „odesilatelů“ zpráv je šíře problematiky poněkud větší a mnohem více různorodá. Každá odesílající infrastruktura je odlišná a bude uplatňovat různá organizační i technická opatření. Jiná pravidla nastaví ISP, který poskytuje pouze čistou konektivitu svým zákazníkům a jejich datový provoz by neměl žádným dalším způsobem ovlivňovat na aplikační vrstvě, jiná pravidla může nastavit obchodní společnost nebo korporace, v jejíž síti se vyskytují pouze zaměstnanci vázaní přesnými pravidly. V jiné situaci bude universita poskytující síťové služby svým studentům, zaměstnancům, ale i hostujícím profesorům, v rámci roamingových dohod studentům jiných universit a například i

účastníkům krátkodobých akcí, jako jsou konference. Specifické postavení mohou mít různé „komunitní“ sítě a velmi specifické postavení budou mít sítě typu hotelové připojení pro hosty, hot-spot sítě v kavárnách, u benzinových pump a podobně.

### 2.3.1. Analýza potřeb a chování uživatelů sítě

Z pohledu ochrany proti spamu, který v chráněné síti vzniká, je vždy třeba provést analýzu chování a potřeb uživatelů sítě. V případě, že síť je teprve budována, je třeba stanovit služby, které svým uživatelům budeme poskytovat a které případně nikoliv. Jedná-li se o síť poskytující konektivitu provozovanou ISP, nejvhodnějším opatřením bude zřejmě pouze přijetí přiměřených pravidel užití, zapovídajících nežádoucí aktivity a hlavně umožňujících aplikaci dostatečně účinných opatření v případě porušení těchto pravidel (zejména možnost dočasného blokování určitého provozu až po možnost vypovězení služeb).

Pokud řešíme problematiku na úrovni sítě koncové, která má vlastní uživatele a těmto uživatelům poskytuje konkrétní aplikační služby, je třeba se zaměřit v analýze na používání konkrétní služby uživateli.

V případě emailu s ohledem na ochranu proti spamu nás bude zajímat zejména:

- Zda odesílají uživatelé poštu s námi spravovanou doménou ze svých MUA pomocí protokolu SMTP, nebo některého proprietárního (Lotus Notes, MS Exchange).
- Zda čtou poštu prostřednictvím POP3/IMAP nebo jiných protokolů.
- Zda potřebují naši uživatelé přístup k elektronické poště i v okamžicích, kdy jsou připojeni mimo domovskou síť.
- Zda v případě přístupu k elektronické poště zvenčí rozlišujeme technologie připojení (například přístup standardními protokoly z jakéhokoliv počítače oproti připojení firemního notebooku do firemní sítě pomocí VPN).
- Zda potřebují uživatelé číst a odesílat ze stanic ve zkoumané síti i poštu s jinou, externí doménou, uloženou na serverech mimo zájmovou síť.
- Zda mají uživatelé vnitřními pravidly sítě, případně jinými pravidly, povoleno číst a odesílat na stanicích v síti „externí“ poštu s cizí doménou odesílatele.

- Zda potřebují v síti SMTP komunikaci nějaká další zařízení a jaká jsou jejich případná technologická omezení.

Na základě zjištěných skutečností bude možné stanovit technická opatření proti riziku šíření spamu z vnitřní sítě a přitom vždy bezpečným způsobem reagovat na zjištěné potřeby uživatelů, případně tyto potřeby organizačním opatřením uživatelům zapovědět.

### 2.3.2. Přiměřená pravidla užití služby

Prvním krokem k ochraně infrastruktury před zneužitím k rozesílání spamu je přijetí přiměřeně nastavených pravidel užití sítě. Obvyklým způsobem v tomto případě je buď ustanovení zakotvené v obchodní smlouvě uzavírané mezi poskytovatelem připojení nebo služby a zákazníkem. Je žádoucí přímo specifikovat způsob řešení spam incidentů, které budou provozovateli hlášeny, a nastavit přípustná technické a organizační opatření. V případě, že jde o koncovou síť, obvykle je vhodnou formou přijímán vnitřní předpis závazný pro uživatele sítě, který zapovídá přímé spamování a v mnoha případech i omezuje kvantitativně množství odesílané elektronické pošty, jež je přípustné. Vzhledem k tomu, že tvrdá čísla jsou v tomto případě problematická, lze se setkat s pravidly typu:

- vyučující může rozeslat email všem svým studentům
- garant předmětu může rozeslat email všem studentům daného předmětu
- student může oslovit své spolužáky v kruhu
- děkan a proděkani fakulty mohou oslovit všechny studenty fakulty
- zpráva zaslaná na více než 100 adres najednou nesmí být větší než xx kB, přílohy je přípustné distribuovat pouze prostřednictvím URL ke stažení

### 2.3.3. Sledování spam reportů a stížností a reakce na ně

V každé síti by měly být zřízeny a sledovány emailové adresy pro hlášení incidentů. Obvykle se jedná o adresy network@domena.tld, abuse@domena.tld a případně postmaster@domena.tld. Zejména první dvě adresy by měly být v každém případě funkční a sledované. Dojde-li například k zařazení některé ze spravovaných IP adres do některého DNSBL, jde o této akci informace na uvedené adresy. Správce pak může reagovat, prozkoumat důvody, které vedly k zařazení na seznam, přijmout potřebná opatření a svojí

reakcí obvykle opět zajistit odstranění ze seznamu dříve, než dojde v větším technickým potížím, případně než dojde k vážnému poškození reputace sítě.

### 2.3.4. Sledování odchozí pošty a nastavení limitů

Jednoduchým technickým opatřením je sledování odchozí pošty a podrobení takové pošty stejnému nebo obdobnému typu kontroly, jako u pošty příchozí. Odchozí poštu lze tedy podrobit například pravidlové analýze a při dosažení vysokého skóre začít poštu z určitého stroje blokovat. Nejjednodušším opatřením je sledování množství odchozí pošty z určité stanice v čase a při překročení nastavené hranice (například 10 emailů za minutu) poštu blokovat a prověřit, zda se nejedná o „zombie“ počítač, ovládnutý útočníkem a zneužitý k rozesílání spamu nebo napadaný virem. V souvislosti s tím je vhodné poskytnout oprávněným uživatelům vhodné nástroje k rozesílání většího počtu zpráv definovaným skupinám, vytvořené tak, aby neúměrně nepřetěžovaly provoz vlastního emailového subsystému.

Jistá opatření je vhodné aplikovat i na odesílání pošty z centrálního mail serveru, například dávkování zpráv pro jednu IP adresu po rozumných kvantech. Může dojít například k situaci, kdy obešleme velkou část uživatelů (např. studentů) vnitřní sítě zprávou a značný podíl těchto uživatelů má nastaveno přesměrování na stejnou freemailovou službu. V takovém případě by došlo k odesílání velkého množství emailů na cílový server freemailové služby najednou a náš server by mohl být vyhodnocen jako spamující a příjem pošty od něj odmítnut nebo omezen. V krajní situaci by toto mohlo vést i ke ztrátě nějakého množství zpráv.

### 2.3.5. Blokování portu 25 a nastavení pravidel pro odesílaní pošty prostřednictvím SMTP

Technické opatření, které je pro ochranu vnitřní sítě obvykle účinné, spočívá v zablokování přístupu na port 25 (protokol SMTP) mimo vnitřní síť na hraničních routerech a zablokování akceptování spojení na port 25 na mailových serverech uvnitř sítě. Centrální mail server určený k odesílání zpráv mimo síť pak následně akceptuje taková spojení pouze z definovaných počítačů (serverů) v síti, určených pro zpracování elektronické pošty. Uživatelé ze svých stanic mohou poštu odesílat pouze pomocí autorizovaného spojení (SMTP-AUTH), s ohledem na bezpečnost optimálně pouze v případě spojení šifrovaného

pomocí SSL. Šifrovaná komunikace protokolu SMTP obvykle probíhá na portu 465, tento tedy musí být povolený pro komunikaci uvnitř sítě mezi jednotlivými segmenty a též zvenčí pro zajištění možnosti přístupu uživatelů k poštovním službám z cizích sítí. Uživatelé zpravidla odesílají poštu přes stejné servery, na kterých poštu přijímají. Pokud umožňujeme zpracovávání i pošty mimo vnitřní doménu (například ve škole), je třeba v případě autorizovaného spojení povolit na kontaktovaných serverech relay, tedy situaci, kdy server přijme a zpracuje poštu i v případě, kdy ani doména odesílatele, ani doména příjemce nejsou vnitřní doménou, pro kterou server poštu zpracovává.

Odesílaní pošty prostřednictvím autentizovaného a šifrovaného spojení je vhodné též používat pro odesílaní pošty s vnitřní doménou při přístupu k mail serveru zvenčí, tedy z IP adres mimo vnitřní síť.

Pro zajištění možnosti práce s elektronickou poštou ze zařízení mimo vnitřní síť a mimo kontrolu správy vnitřní sítě (například počítače v internetových kavárnách) je vhodné zpřístupnit WWW rozhraní pro práci s poštou.

V případě, že pod správu vnitřní sítě spadá i určitý rozsah nebo určité rozsahy IP adres, které nelze omezit buď z technických nebo organizačních důvodů striktními pravidly na hraničních routerech (například se jedná o dislokované pracoviště s problematicky zajištěnou konektivitou nebo jde o WiFi síť určenou pro provoz návštěvníků konference) a nelze se takové situaci vyhnout, je žádoucí příslušné rozsahy IP adres přímo zadat na vybrané DNSBL a tím alespoň částečně zamezit přijímaní pošty z takových rozsahů.

Pokud můžeme definovat striktní pravidla pro poštovní provoz emailových adres s kontrolovanou doménou, je poměrně vhodné využít všech možností specifikace odesílajících serverů elektronické pošty. Je možné definovat v TXT záznamu naší domény v DNS systému platný SPF záznam, který přesně konkretizuje, které servery mohou odesílat poštu s adresou odesilatele z kontrolované domény. Poměrně obvyklým případem je situace, kdy pošta je odesílána přes stejné servery, přes které je i přijímána, lze tedy SPF záznam omezit na MX záznamy v DNS, případně uvést přímo IP adresy strojů, které používáme pro odesílání pošty, nejsou-li stejné jako MX servery. Na takovéto omezení je důležité upozornit všechny uživatele a zejména zdůraznit rizika problémů v případě přeposílání pošty s chráněnou doménou z jiných sítí, případně rizika služeb, jako jsou různé servery WWW služby odesílající elektronické pohlednice a podobně.

### 2.3.6. Shrnutí technik ochrany z pohledu odesílající infrastruktury

Je zřejmé, že ochranu spravované infrastruktury před zneužitím spamery je nutno chápat jakou součást obecné bezpečnostní politiky uplatňované uvnitř sítě. Celý proces je třeba stavět na přesně definovaných administrativních opatřeních a používání určených sankcí v případě jejich porušení. Pro vlastní ochranu je následně uplatňován princip technické ochrany koncových stanic před škodlivým kódem a zabezpečení stanic před nežádoucím chováním jejich uživatelů. Pro zamezení masivního rozesílání mailů při případném nezachyceném napadení jsou přijímána pravidla znemožňující přímé anonymní odesílaní zpráv ze stanic a usměrňování provozu tak, aby data procházející přes hranici sítě mohla být v případě potřeby kontrolována v souladu s bezpečnostní politikou.

Jako další prostředek, bohužel s menší a předem nedefinovatelnou efektivitou, lze vidět použití technik, které okolí dostatečně věrohodně definují - kdo, za jakých okolností a odkud může odesílat zprávy s naší doménou. Lze definovat SPF záznamy, zvážit hodnocení sítě v reputačním programu a podobně. Potenciální ostatní příjemce spamu ale nelze přimět k respektování takových opatření, a proto je jejich efekt nejistý. U opatření uplatňovaných na straně odesílající infrastruktury je vždy nutné pečlivě zvážit, jaká omezení jejich nasazení přinese, jaké náklady vygeneruje a zda jejich přínos je tomu odpovídající.

# 3. KOMENTÁŘOVÝ SPAM

Cíle a důvody komentářového spamu jsou podobné cílům, které má spam v elektronické poště. Jejich podstata spočívá v rozšíření informace o určitém produktu nebo službě, obvykle ve spojení s příslušnými URL, a zásahu co největšího množství čtenářů. Technika komentářového spamu je využívána proti diskusním fórům či diskusním serverům na Internetu nebo proti diskusím u článků na zpravodajských či zájmových serverech. Vzhledem k tomu, že zde je spam umísťován do dokumentů dostupných prostřednictvím protokolu HTTP, jde často i o ne zrovna čistou praktiku SEO, provozovanou s cílem zvýšit PageRank či obdobný hodnotící index vyhledávacích serverů pomocí umístění zpětných odkazů na servery, které mají zajímavou reputaci ve výpočtu takového indexu. Výsledkem je zlepšení pozice propagovaných stránek ve zobrazovaných výsledcích na nějaké konkrétní vyhledávané heslo.

## 3.1. Metody obrany proti komentářovému spamu

Obrana proti komentářovému spamu je obvykle založena na nějaké modifikaci Touringova testu; jde tedy o snahu automatizovaně rozpoznat, zda na straně odesílatele komentáře je živý člověk, nebo stroj.

### 3.1.1. Registrace přispěvatelů

První a obvykle nejjednodušší formou obrany proti komentářovému spamu zadávanému automaty je zavedení povinné registrace. Před odesláním komentáře je pak zpravidla vyžadována autorizace osoby. Aby registrace nebyla proveditelná též automaticky, často se používá ověření pomocí zaslání unikátní zprávy na uživatelem zadanou emailovou adresu. Následné zadání obdrženého kódu (odkazu) je nezbytné pro úspěšné dokončení registrace.

### 3.1.2. Náhodná změna jmen polí v HTML formulářích

Tato metoda je založena na dynamickém vytváření HTML formuláře tak, aby nebylo předem jasné, jaká jména polí se v něm vyskytují. Odesílací robot má pak v takovém případě značně ztíženou pozici. Metodu lze zefektivnit například tím, že po prvním odeslání je příspěvek zobrazen uživateli znovu v editovatelné podobě ke korektuře, ale s náhodně vygenerovanými jmény polí a jejich správnost je následně při druhém odeslání zkontrolována.

### 3.1.3. CAPTCHA

Captcha je akronym „Completely Automated Public Turing test to tell Computers and Humans Apart“ [9]. Jde o techniku, kdy je ve formuláři pro zapsání příspěvku zobrazován obrázek s náhodným textem, který je třeba přepsat do příslušného pole. Obraz je vytvářen tak, aby byl pro člověka snadno rozeznatelný, ale pro počítač pomocí technologie OCR naopak nerozeznatelný vůbec. Technika má několik nevýhod. Jednak je proti zásadám přístupnosti webu, neboť brání v užití stránek zrakově postiženým lidem, a jednak trpí stále větší nečitelností obrázků danou tím, jak se zlepšují možnosti techniky OCR.

### 3.1.4. Umělá inteligence

Testy jsou většinou založeny na snaze „vyzkoušet“ zadavatele pomocí jednoduchého testu inteligence. Na serverech psaných česky (nebo jiným, ne příliš rozšířeným jazykem) obvykle stačí zadat nějakou jednoducho otázku, na kterou se všeobecně předpokládá, že přispěvatel zná odpověď, a tuto odpověď pak zkontrolovat. Stačí tedy otázka „Kolik je jedna plus jedna?“ a většina robotů je odfiltrována. Vyskytují se dokonce tak extrémní implementace této techniky, jako je popsaná v článku „Konečné řešení v boji proti spamu v diskusních fórech“ [10], které otázku nebo pokyn kódují do obrázků deformovaného podobně jako pomocí Captcha, do něhož ještě vkládají například překlepy, aby analýza nebyla jednoduchá ani prostředky jednoduché umělé inteligence.

# 4. DALŠÍ PŘÍKLADY SPAMU

## 4.1. Usenet News spam

Usenet News spam je v dnešní době již nedůležitým problémem - potřeba boje s tímto druhem spamu vymizela s postupným ústupem užívání systému Usenet News. Svoji podstatou byl Usenet spam podobný spamu emailovému, k efektivnímu šíření využíval techniku Usenet News serverů, které se v hierarchické struktuře zrcadlily po celém světě. K distribuci zprávy tedy stačilo doručit ji pouze k nejbližšímu Usenet News serveru a technologie sama se postarala o další replikaci spamu.

## 4.2. Spam over Instant messaging (SPIM)

Se vzestupem využívání systémů IM pro online komunikaci vystupuje lehce na povrch i problém spamu v tomto prostředí. Vzhledem k tomu, že IM systémů existuje relativně velké množství, jejich komunikační protokol není vždy zcela otevřený, jednotlivé systémy mezi sebou nejsou provázány a obvykle přímo obsahují metody autorizace uživatelů navzájem, není problematika SPIMu příliš obsáhlá a většinou nevyžaduje konkrétní řešení kvůli například zahlcení sítě. Vzhledem k tomu, že jsou známy případy šíření počítačových virů prostřednictvím IM, je vhodné v případě požadavků na bezpečnost bránit používání globálního IM systému a pouze například v rámci lokální sítě zavést oddělený ostrůvek IM.

## 4.3. Spam over IP telephony (SPIT)

Vzhledem k tomu, že IP telefonie není zatím příliš rozšířena a masově používána ve formě takové, kdy by přímo koncová zařízení byla připojena do globální sítě, není problém tzv. SPITu zatím příliš palčivý. Lze však očekávat pokusy o uplatnění i této techniky pro telemarketing a podobné účely. Požadavky na technické vybavení „útočníka“ jsou ale v tomto případě řádově vyšší než u běžného spamu šířeného elektronickou poštou.

# 5. SHRNUTÍ TEORETICKÉ ČÁSTI

Teoretická část práce se zaměřuje zejména na definici spamu v prostředí určeném k přenosu zpráv. Popisuje charakteristiky a příklady spamu primárně v systému elektronické pošty založené na protokolu SMTP, částečně je pozornost věnována i spamu komentářovému, vyskytujícímu se na WWW stránkách s diskusními fóry nebo na webech s možností komentovat nějaký subjekt (článek, výrobek, akci, …). Okrajově je zmíněna problematika spamu v instant messaging systémech (ICQ, Yahoo! Messenger, …) a prostředí IP telefonie.

Těžiště teoretické části práce se zaměřuje na principy a možnosti ochrany proti spamu v elektronické poště. Postupně je ochrana diskutována z pohledu individuálního příjemce emailů, jeho možného chování a ochranných mechanismů, omezujících obtěžování spamem. Zevrubně je diskutována automatizovaná ochrana proti spamu z pohledu infrastruktury, stojící na straně příjemců elektronické pošty, protože tento pohled je ve většině případů zajímavější a nejpalčivější. Velká většina organizací se snaží omezit obtěžování a zdržování svých zaměstnanců spamem i snížit technické nároky na provoz elektronické pošty, které mohou kvůli obrovskému objemu spamu narůst do neudržitelných rozměrů.

Pozornost je věnována i problematice ochrany koncové sítě před zneužitím k rozesílání spamu, případně jiného škodlivého nebo obtěžujícího kódu.

# II. PRAKTICKÁ ČÁST

# 6. DEFINICE VÝCHOZÍCH PŘEDPOKLADŮ

V praktické části práce se budeme zabývat návrhem antispamových opatření, uplatňovaných v modelové organizaci s většími objemovými nároky na emailový provoz. Aby opatření byla pokud možno účinná a komplexní, návrh bude sestávat ze schematického technického řešení emailové infrastruktury, zavedení příslušných opatření a ukázky jejich konkrétní realizace v nastíněné infrastruktuře. Návrh bude též definovat základní pravidla a organizační opatření určená k odvrácení rizika stát se zdrojem spamu a technická opatření ke zmírnění takového rizika vedoucí.

## 6.1. Modelová organizace

Aby návrh opatření a technických prostředků byl pokud možno universální, dal se aplikovat i na menší organizace, případně jej bylo možné považovat za částečně modulární, a byl investičně pokud možno nenáročný, předpokládejme že navrhujeme systém pro organizaci umožňující využívání open source softwarových produktů a pracující v přibližně následujících podmínkách:

- Organizace provozuje IT infrastrukturu založenou na běžných prostředcích výpočetní techniky. Vnitřní komunikace je založena na protokolu TCP/IP. Operační systémy a HW prostředky koncových zařízení jsou různorodé a nesmí být omezující pro navržené řešení.
- Infrastruktura je rozsáhlá, rozdělená do většího počtu lokalit, adresní prostor zapojených zařízení je celistvý a uzavřený. Hraniční router nebo routery jsou pod centrální správou a kontrolou. V síti jsou zapojená i zařízení, která nevlastní provozovatel infrastruktury a nemá nad nimi přímou kontrolu (například soukromé počítače).
- V síti pracuje velké množství (řádově tisíce) uživatelů, pro které jsou zajišťovány základní služby. Předpokládejme, že máme dvě skupiny uživatelů s různými nároky na dostupnost a funkcionalitu služby, které jsou řešené různými prostředky (například zaměstnanci, požadující vysokou spolehlivost služby a dodatečné funkce systému – například groupwarové, time management a podobně, ve skupině jedné a studenti, kterých je násobně více, ale vyžadující pouze funkcionalitu zasílání zpráv). Obě skupiny jsou navíc nejednotně organizované.

Management identit (uživatelských jmen, oprávnění) je vyřešen centrálně a potřebná data jsou k dispozici.

- Z chování a zkušeností víme, že přinejmenším jedna skupina vyžaduje ze stanic ve vnitřní síti organizace i přístup k externí elektronické poště, včetně možnosti odesílat zprávy s cizí doménou, a tomuto požadavku chceme vyhovět.
- Emailové systémy, které obsluhují přímo poštovní schránky uživatelů v rámci jejich požadavků a potřeb na funkcionalitu, nejsou řešeny; předpokládejme že ve vyhovující formě existují a jsou schopny navenek komunikovat protokolem SMTP.

Vynecháme-li specifický svět ISP, představuje tento navržený model asi nejkomplexnější požadavky na antispamovou ochranu sítě nebo služby, s jistými úpravami použitelný například i pro provozovatele freemailové aplikace.

## 6.2. Technická východiska

Technicky stojíme před problémem volby typu centrálního mailserveru (resp. mailserverů), jeho nastavení a umístění v síti a nastavení pravidel provozu na hraničních routerech sítě s ohledem na mailový provoz. Předpokládáme, že síťová problematika infrastruktury je vyřešená nebo snadno řešitelná a dodavatel síťových služeb na případné požadavky reaguje. Veškeré prostředky nutné pro technický provoz sítě (například DNS, přidělování IP adres, roubovací pravidla uvnitř sítě, …) jsou plně pod naší kontrolou.

# 7. NÁVRH ŘEŠENÍ

## 7.1. Organizační opatření a omezení mailového provozu v rámci vnitřní sítě

V předpokladech jsme definovali, že navrhujeme antispamovou ochranu v rozsáhlejší síti, ve které se vyskytují různá koncová zařízení, a platí situace, že ne všechna zařízení do sítě připojená jsou pod kontrolou správce infrastruktury. Zároveň je třeba, aby uživatelé měli možnost z takovýchto zařízení odesílat i poštu s cizí doménou.

Je zřejmé, že je třeba zabránit živelnému odesílání emailových zpráv z koncových stanic uvnitř sítě přímo na externí MTA mimo naši síť. Zvolené MTA uvnitř sítě budou sloužit jako relay (předávací servery), na kterých bude možné provádět kontroly a vyhodnocovat oprávněnost odeslání pošty. Protože mnoho virů i trojských koňů užívaných k odesílání spamu dokáže z nastavení obvyklých MUA na stanicích se běžně vyskytujících zjistit, přes který MTA je třeba poštu předávat, pouhé blokování portu 25 na hraničních routerech je opatření první, nutné, avšak nikoliv dostačující. Pro odesílání pošty z koncových stanic pomocí protokolu SMTP je nutné zavést autorizované šifrované spojení. Mezi první opatření tedy bude patřit organizační směrnice, určující, že veškeré stanice v síti mohou odesílat poštu pouze prostřednictvím definovaných MTA a spojení je nutné navazovat autorizované a šifrované pomocí SSL. Předávací servery mohou být stejné, jaké slouží k uchovávání poštovních schránek uživatelů, nebo to mohou být i jiné servery, které si dokáží ověřit autentizační údaje. V souvislosti s tím je nutné změnit konfiguraci koncových mailserverů tak, aby spojení na port 25 přijímaly pouze z definovaných centrálních mailserverů a nikoliv od zařízení ve vnitřní síti ani od zařízení mimo tuto síť. Toto opatření je možné zdvojit o nastavení příslušných pravidel na routerech uvnitř sítě.

V druhém kroku je nutno zvolit, zda si přejeme odesílání zpráv s doménou naší sítě pouze prostřednictvím námi kontrolovaných MTA. Pokud ano, nastavíme vhodný SPF záznam do DNS systému a o této skutečnosti informujeme uživatele s upozorněním na možná rizika tohoto opatření (zejména přeposílání zpráv). Je nutné v první řadě seznámit uživatele s tím, že emaily s vnitřní adresou je vždy nutné odesílat prostřednictvím určených MTA uvnitř sítě, připojovat se na ně stejným způsobem jako ze stanic uvnitř sítě a v případě, že toto není možné, použít web rozhraní k vnitřnímu systému elektronické pošty.

## 7.2. Volba mailserveru a jeho umístění

### 7.2.1. Umístění mailserverů

Vycházíme z definovaného předpokladu, který říká, že obsluhujeme dvě skupiny uživatelů s různými nároky na poskytnutou službu. Tyto nároky jsou v rámci existujícího řešení pokryty uspokojivě a je pouze potřeba zajistit ochranu proti spamu v mailu přicházejícím zvenčí. Pro potřeby práce nám postačí informace, že koncové mailservery určené k obsluze poštovních schránek uživatelů jsou v síti umístěny v zóně definované bezpečnostní politikou a je s nimi možné komunikovat prostřednictvím protokolu SMTP. Potřebujeme pouze na tyto servery zprávy předávat a od nich zprávy v případě potřeby přebírat.

Pro potřeby komunikace s vnějším světem zvolíme nově centrální mailserver, přes který bude probíhat veškerý příchozí provoz. V případě, kdy to je žádoucí, můžeme provozovat i záložní centrální mailserver, pro zvýšení odolnosti řešení proti výpadkům a zajištění zachování zpráv i při déletrvajícím výpadku spojení mezi různými částmi internetu můžeme tento záložní centrální MTA umístit u jiného poskytovatele připojení. V takovém případě by kritériem mělo být zejména to, aby záložní MTA byl umístěn v jiném autonomním systému, než z jakého získáváme konektivitu pro naši síť, a ve kterém je tedy umístěn primární MTA. Pro umístění obou serverů z hlediska síťového provozu musí platit, že musí být dostupné prostřednictvím protokolu SMTP z celého internetu s výjimkou vnitřní sítě.

IP adresy těchto serverů následně nastavíme jako MX záznamy v DNS s určenou prioritou. V případě, že budou používány i pro odesílání pošty z vnitřní sítě do Internetu a přejeme si, aby pošta s adresou naší domény byla odesílána pouze z námi kontrolovaných strojů, nastavíme v DNS též SPF záznam typu:

```
nasedomena.tld. IN TXT "v=spf1  mx -all"
```

### 7.2.2. Volba SW pro centrální mailserver

V dalším kroku připravovaných opatření je nutné zvolit vhodný SW, který bude zajišťovat provoz centrálních MTA určených pro předávání zpráv mezi vnitřní sítí a okolním světem a na této pozici bude zajišťovat automatizovanou ochranu proti spamu zasílanému do vnitřní sítě. Tímto se stane součástí bezpečnostních opatření chránících síť a vnitřní infrastrukturu. Mezi další úlohy z pohledu bezpečnosti bude s největší pravděpodobností

patřit i ochrana proti virům a dalšímu škodlivému kódu, který může vnitřní síť infiltrovat prostřednictvím elektronické pošty.

Vhodný SW je možné vybírat dle mnoha kritérií, pro potřebu této práce volme kritéria následující, korespondující s obvyklými požadavky reálného provozu:

1. Řešení musí být dostatečně robustní a ověřené. Důvodem je rozsah služby, která má být zajišťována, a požadavek na její spolehlivost a dostatečný výkon. U rozsáhlých infrastruktur obvykle nelze v ostrém provozu experimentovat s prostředky mírně obskurními, i když nadějnými. Současně musí být přiměřeně zajištěna kontinuita produktu.
2. Nasazený SW musí být ekonomicky co nejméně náročný. Nejedná se pouze o pořizovací cenu, ale o TCO, kam lze zahrnout i případnou cenu za licence ke každému obsluhovanému mailboxu, respektive emailové adrese nebo sadě adres, i cenu za roční nebo časovou podporu běžně požadovanou u serverových systémů a též cenu práce s údržbou příslušného systému.
3. Systém musí podporovat požadované antispamové a bezpečnostní techniky, které chceme nasazovat, optimálně musí poskytovat univerzální rozhraní pro implementaci mezizpracování emailů v různých stadiích uskutečňovaného přenosu.

Uvedeným kritériím vyhovuje systém pro přenos elektronické pošty postfix [16]. Vzhledem k předloženým kritériím lze postfix vyhodnotit následovně:

Ad 1) Postfix je opensource projekt, který existuje a úspěšně se rozvíjí již několik let, a za dobu vývoje prošel několika generacemi. Vznikl se záměrem zajistit robustní, snadno konfigurovatelný a bezpečný SW určený pro roli MTA, který by dokázal nahradit v dřívějších dobách téměř výhradně používaný program sendmail a vyvarovat se problémů tohoto SW. Postfix je v současné době používán mnoha organizacemi po celém světě a využívají jej i sítě s velmi vysokým mailovým provozem, jako například univerzity. Vzhledem k jeho vývoji jako opensource je kontinuita zajištěna přiměřeně.

Ad 2) Postfix je opensource projekt, za jehož použití není nutné platit žádný licenční poplatek. Postfix pracuje pod většinou v současné době používaných Unix-like operačních systémů, a lze tedy využít již existující investice v případě, že v organizaci nějaký Unix-like operační systém provozován je. Pokud není, lze postfix nasadit pod všemi běžně používanými distribucemi operačního systému Linux. Pořizovací náklady

tedy v takovém případě odpovídají pouze přiměřeně dimenzovanému HW. Použití systému postfix může generovat jisté náklady na konfiguraci a údržbu provozu v případě, že organizace nemá k dispozici IT pracovníka se zkušeností s provozem Unix-like operačních systémů a MTA.

Ad 3) Postfix je vybaven rozhraním umožňujícím zpracování a kontrolu informací v průběhu SMTP spojení a v průběhu práce s přijatou zprávou v době, kdy je uchovávána ve frontě zpráv určené k doručení. Je k dispozici velké množství aplikací provádějících různá bezpečnostní, antivirová, kontrolní a jiná opatření pro zajištění požadovaného stavu poštovního provozu. Implementace je provedena pomocí možnosti vyhodnotit SMTP relaci ve fázích HELO, CLIENT, SENDER, RECIPIENT a DATA, tedy v okamžicích, kdy došlo k provedení příslušných příkazů SMTP komunikace. Do generování návratového kódu SMTP komunikace je možné zapojit externí programy pomocí „Postfix SMTP Access Policy Delegation“. V určených okamžicích může být vyhodnocení předáno externím programům, které mohou v návratovém kódu poslat neutrální odpověď, nebo požadavek zamítnout se zvoleným chybovým návratovým kódem.

Externí programy mají k dispozici následující údaje dle verze programu postfix:

```
Postfix version 2.1 and later:
request=smtpd_access_policy
protocol_state=RCPT
protocol_name=SMTP
helo_name=some.domain.tld
queue_id=8045F2AB23
sender=foo@bar.tld
recipient=bar@foo.tld
recipient_count=0
client_address=1.2.3.4
client_name=another.domain.tld
reverse_client_name=another.domain.tld
instance=123.456.7
Postfix version 2.2 and later:
sasl_method=plain
sasl_username=you
sasl_sender=
size=12345
ccert_subject=solaris9.porcupine.org
ccert_issuer=Wietse+20Venema
ccert_fingerprint=C2:9D:F4:87:71:73:73:D9:18:E7:C2:F3:C1:DA:6E:04
Postfix version 2.3 and later:
encryption_protocol=TLSv1/SSLv3
encryption_cipher=DHE-RSA-AES256-SHA
encryption_keysize=256
etrn_domain=
[empty line]
```

## 7.3. Navržené automatizované ochranné techniky

V teoretické části práce byly popsány automatizované techniky ochrany infrastruktury maily přijímající. Nynějším úkolem je zvolit a odůvodnit vhodnou kombinaci použitých technik, které povedou k dostatečně efektivnímu filtrování spamu za přijatelných nákladů a se spolehlivostí a přesností dostatečnou pro uživatele.. Je zřejmé, že hodnocení je již předem odsouzeno k jisté neurčitosti, neboť přijatelné náklady a dostatečná spolehlivost a přesnost budou v různých situacích různé.

Přihlédneme-li k výhodám a nevýhodám, případně dalším vlastnostem jednotlivých metod popsaných v teoretické části, je zřejmé, že prvním opatřením v řadě by měly být metody, které probíhají ještě před vlastním přenesením dat. Ušetříme tím přenosové pásmo připojení a též značnou část strojového času potřebného pro následnou analýzu dat.

Z technik použitelných před přenosem zpráv máme možnost testovat odesílající server na existenci záznamů v blacklistech, korektnost provedení příkazu HELO/EHLO a můžeme použít greylisting. Jak uvádí Kára v [4] a [5], používání blacklistů jako striktních (tvrdých) pravidel přinese pravděpodobně přes svoji jednoduchost více problémů než užitku a na tomto místě není nejšťastnější jej použít. Vyhodnocování implementace HELO/EHLO příkazu možné je. Za poměrně výhodnou metodu lze považovat greylisting. Sice přinese nezanedbatelný požadavek na strojový čas, neboť z principu metody bude nutné zpracovávat charakteristické trojice (IP adresa odesílajícího serveru/mail adresa odesilatele/mail adresa příjemce) velkého množství zpráv, avšak pravděpodobně dokáže odmítnout velké množství spamu a virových zpráv bez přenesení a následné analýzy dat. Pro odlehčení náročnosti na strojový čas je vhodné greylisting kombinovat s whitelistem, obsahujícím IP adresy serverů, se kterými často komunikujeme, a dále takové adresy, které je globálně vhodné whitelistovat (například odesílací farmu Gmailu).

Jisté další odlehčení nároku na výpočetní výkon lze získat předřazením triviální kontroly, která dokáže část zpráv odfiltrovat účinně. Za takovou kontrolu lze považovat test SPF záznamů odesílající domény. Vlastní kontrola a vyhodnocení jsou velmi jednoduché, data jsou získávána prostřednictvím systému DNS, a lze takto odmítnout určitou část zpráv bez dalšího uchovávání jakýchkoliv záznamů potřebných pro provoz metody.

Protože je pravděpodobné, že část spamu přes takovéto kontroly projde, na konec řetězce vyhodnocování je vhodné zapojit některou metodu založenou na analýze obsahu. Takto

vyhodnocené zprávy je však nejrozumnější pouze označit a jejich další osud již nechat na koncovém uživateli.

# 8. VOLBA HODNOTÍCÍCH RUTIN A KONFIGURACE MTA

## 8.1. Dostupné Policy access servery

V dokumentaci a na domovské stránce projektu Postfix [16] najdeme seznam existujících dostupných policy serverů, které lze přímo použít pro naše záměry. Rovněž distribuce postfixu obsahuje příklad greylist policy serveru napsaného v jazyce Perl a pracujícího s jednoduchými Unix db soubory.

Seznam je dostupný na adrese http://www.postfix.org/addon.html. K dispozici jsou následující servery:

- *apolicy*: server napsaný v jazyce Python umožňující sestavování pravidel stylu ACL a jejich vyhodnocování pomocí regulárních výrazů
- *ppolicy*: server napsaný v jazyce Python s modulární architekturou. Implementuje moduly pro geografické určení IP adres, vyhledávání IP adresa v zadaných blacklistech a případný výpočet skóre na základě většího množství blacklistů, nabízí modul pro greylisting, i kontrolu SPF. Architektura umožňuje vytvářet další moduly dle potřeby.
- *gld*: greylist policy server využívající pro ukládání dat o tripletech databázi MySQL nebo PostgreSQL. Psán je v jazyce C.
- *SQLgrey*: greylist policy server psaný v jazyce Perl; pro ukládání dat umí využívat databáze MySQL, PostgreSQL nebo SQLite. Umožňuje sofistikovanější verzi whitelistování a v určitých případech dokáže na whitelist přidávat automaticky servery pozitivně používající SRS.
- *gps*: greylist policy server napsaný v jazyce C++; pro ukládání dat umí využívat databáze MySQL, PostgreSQL nebo SQLite. Propracovává několik metod pro vlastní greylistování a umožňuje efektivně whitelistovat dle DNS nebo IP adres.
- *Postgrey*: greylist policy server využívající databázi PostgreSQL.
- *policyd*: v jazyce C napsaný server kombinující několik metod vyhodnocování emailů (white/grey/blacklisting, HELO kontroly, omezení kvót a podobně).

- *policyd-weight*: policy server umožňující vyhodnocování DNSBL, HELO, FROM a klientských IP adres s možností definovat váhy a výsledek určovat na základě dosaženého skóre.
- *tumgreyspf*: greylisting a SPF policy server využívající pro ukládání dat pouze vlastní filesystem.

Přímo v distribuci postfixu je doporučen policy server pro kontrolu SPF záznamů. Na internetu lze nalézt případné další policy servery přizpůsobitelné nebo navržené dle očekávané funkce a vhodné pro požadované řešení.

## 8.2. Použité Policy access servery

V předchozí kapitole jsme navrhli použití kontrol SPF a greylistingu s možností whitelistování odesílatelů zpráv dle potřeby.

Pro SPF kontrolu je nejjednodušší použít přímo policy access server doporučený v distribuci postfixu. Policy server je dostupný na http://www.openspf.org/Software [17].

Pro greylistování lze použít větší množství policy access serverů. Pro potřeby této práce zvolíme SQLgrey z následujících důvodů:

- je napsaný v jazyce Perl, který zachovává programy v jejich původní zdrojové podobě, a jejich případná úprava a oprava je tedy jednoduchá, přitom ale při startu dochází k online kompilaci kódu a vlastní provoz serveru již není interpretovaný s příslušnými nevýhodami
- je aktuálně stále vyvíjený
- umožňuje použít různé databáze dle možností a potřeby
- obsahuje automaticky udržovaný whitelist zdrojů zpráv, u kterých je tato potřeba dána technologicky (nevhodné konfigurace, odesílací pooly velkých služeb a podobně)
- umožňuje snadno udržovat lokální whitelist budovaný dle potřeby
- umožňuje zvolit greylistovací metodu (striktně dle IP adresy nebo pouze dle části IP adresy odesílatele) s ohledem na generovanou zátěž serveru

- na základě cílových emailových adres umožňuje volit optin nebo optout greylistování – je tedy možné volit greylisting pouze pro některé cílové emailové adresy nebo některé naopak z greylistingu vyjmout. V případě obsluhy většího množství domén, pro které server přijímá elektronickou poštu, je možné pro různé domény tyto metody kombinovat dle potřeby.

## 8.3. Konfigurace postfixu

Pro použití policy access serverů v postfixu je třeba provést několik kroků v konfiguraci serveru:

- nainstalovat do systému příslušný policy access server
- do konfiguračního souboru master.cf postfixu (obvykle se nachází v adresáři /tec/postfix) v seznamu démonů doplnit zvolený policy server dle potřeby v případě, že s ním hlavní proces bude komunikovat prostřednictvím unix socketů. Bude-li volena komunikace prostřednictvím TCP na zvoleném portu, lze policy server zapsat přímo do konfigurace služby.
- do konfiguračního souboru main.cf postfixu (obvykle se nachází v adresáři /tec/postfix) doplnit policy službu do příslušné kontrolní direktivy

Konfigurační soubory mohou vypadat například následovně:

```
Master.cf:

#==========================================================================
# service type  private unpriv  chroot  wakeup  maxproc command + args
#               (yes)   (yes)   (yes)   (never) (100)
#
#==========================================================================
greylist        unix    -       n       n       -       -       spawn
   user=robot  argv=/usr/bin/perl5 /usr/local/libexec/greylistmysql.pl
spf     unix    -       n       n       -       30      spawn
   user=robot  argv=/usr/bin/perl5 /usr/bin/spf.pl


Main.cf:
smtpd_recipient_restrictions =  permit_mynetworks,
      reject_unauth_destination,
      check_policy_service unix:private/spf,
      check_policy_service unix:private/greylist
```

Tučně označené části označují konkrétní údaje týkající se policy access serverů.V tomto ukázkovém případě jsou použity konfigurace pro SPF kontrolu a greylist kontrolu pomocí

velmi jednoduchého serveru bez možností sofistikovanější konfigurace a whitelistingu napsaného v perlu (greylistmysql.pl). S oběma servery je v tomto případě komunikováno prostřednictvím unix socketů.

### 8.3.1. Zprovoznění SPF

Policy server postfix-policyd-spf-perl lze stáhnout na domovské stránce projektu Sender Policy Framework [17]. Instalace policy serveru je jednoduchá: po rozbalení distribučního balíčku je třeba zkopírovat soubor `postfix-policyd-spf-perl` do vhodného adresáře na operačním systému, ze kterého je možné spouštět programy. V dokumentaci je doporučován adresář `/usr/local/lib,` ale je možné zvolit i jakýkoliv jiný.

Po provoz postfix-policyd-spf-perl je třeba mít nainstalován interpret jazyka perl a následující podpůrné balíčky:

- Perl 5.6
- version
- NetAddr-IP 4
- Mail-SPF

V případě, že podmínka není splněna, je třeba software doplnit; vlastní proces závisí na operačním systému a případně verzi a distribuci použitého systému.

Upravíme konfigurační soubor master.cf následovně:

```
spf unix  -       n       n       -       -       spawn
         user=nobody argv=/usr/bin/perl /usr/local/lib/policyd-spf-perl
```

a konfigurační soubor main.cf takto:

```
smtpd_recipient_restrictions =

            ...
            reject_unauth_destination
            check_policy_service unix:private/spf
            ...
```

Následuje restart serveru postfix.

### 8.3.2. Zprovoznění SQLGrey

Policy server SQLgrey je dostupný na domovské stránce projektu SQLgrey hostované na SourceForge.net [18]. Pro provoz je nutné mít k dispozici následující SW:

- Perl
- Net::Server
- IO::Multiplex
- perl-DBI
- Postfix 2.1 (nebo novější)

A je vhodné mít k dispozici některou z podporovaných SQL databází, tedy MySQL nebo PostgreSQL.

Vlastní zprovoznění proběhne v několika krocích:

- Vytvoření adresáře s konfigurací serveru (doporučován adresář `/etc/sqlgrey`).
- Vytvoření uživatele, pod kterým bude server spouštěn (doporučeno sqlgrey).
- Zkopírování konfiguračních souborů z distribuce (z adresáře `etc`).
- Úprava souboru `/etc/sqlgrey/sqlgrey.conf` (zejména nastavení zvolené databáze).
- Tvorba databáze, nastavení práv. Databázové tabulky se vytvoří při prvním spuštění. Struktura použité databáze je uvedena v příloze 1.
- Konfigurace postfixu.

Policy server SLQgrey komunikuje s postfixem prostřednictvím TCP/IP. Není nutné v konfiguračním souboru master.cf definovat nový typ serveru/služby, ale je možné zapsat kontrolu přímo do souboru main.cf. Konfigurační soubor bude vypadat následovně:

```
smtpd_recipient_restrictions =
                ...
                reject_unauth_destination
                check_policy_service inet:127.0.0.1:2501
```

### 8.3.3. Konfigurace SPF a SLQgrey společně

Policy serverů můžeme samozřejmě používat víc a zprávy lze vyhodnocovat postupně proti několika různým policy serverům. Každé vyhodnocení může obecně končit třemi stavy – OK (akceptujeme email) / dunno (nevím, nedělám nic) / reject (odmítnu zprávu s určeným chybovým kódem). Celá sestava se chová jako posloupnost pravidel, ve které se uplatní první pravidlo v pořadí, které vrátí výsledek OK nebo reject.

V případě kombinace SPF a greylistingu je zjevně výhodné první provádět jednoduchý a rychlý test na SPF a v druhém kroku nasadit greylisting. Pořadí vyhodnocování závisí na pořadí zápisu policy serverů v souboru main.cf. V našem konkrétním případě bude tedy příslušná část main.cf vypadat následovně:

```
smtpd_recipient_restrictions =
            ...
            reject_unauth_destination
            check_policy_service unix:private/spf
            check_policy_service inet:127.0.0.1:2501
            ...
```

Tímto krokem je úspěšně nastavena základní antispamová ochrana, které může ve stejné podobě fungovat jak na centrálním mailserveru, který přijaté zprávy distribuuje na další servery obsluhující konkrétní uživatelské schránky, tak může zcela stejně fungovat na koncovém mailserveru, obsluhujícím uživatelské schránky přímo.

# 9. VYHODNOCENÍ OCHRANY SPF A GREYLISTING

Zhodnocení navrhované ochrany lze provést pouze empiricky na experimentálních nebo standardně provozovaných instalacích po určité době provozu.

Zhodnocení navrženého řešení vychází ze zkušeností autora s provozem uvedených ochran na serveru `io.pinknet.cz`, užívaném pro hostování WWW projektů a nabízejícím svým uživatelům i obsluhu emailových schránek, včetně schránek virtuálních, souvisejících s provozovanými WWW projekty. Dalším zdrojem informací je provoz centrálního mailserveru Vysoké školy ekonomické v Praze.

## 9.1. Vyhodnocení ochrany metodou kontroly SPF záznamu

Pro hodnocení účinnosti kontroly SPF jsou použita data logů mailového provozu na serveru `io.pinknet.cz` za období 15.4.2007 až 21.5.2007. V každém dni byl sledován počet výskytů vyhodnocení DUNNO (neutrální výsledek, zpráva je podrobena dalšímu testování) pomocí kontroly SPF a vyhodnocení REJECT (záporný výsledek, tzn. doména odesilatele deklaruje pravidla, odkud mohou být odesílány zprávy s uvedenou doménou odesilatele, a zpráva deklaraci neodpovídá). Pro každý den je uveden procentuální podíl odmítnutých a neutrálně vyhodnocených zpráv.

| Datum | DUNNO | REJECT | % |
|---|---|---|---|
| 15.4.2007 | 4992 | 243 | 4,87 |
| 16.4.2007 | 10363 | 200 | 1,93 |
| 17.4.2007 | 9125 | 175 | 1,92 |
| 18.4.2007 | 9291 | 218 | 2,35 |
| 19.4.2007 | 9751 | 802 | 8,22 |
| 20.4.2007 | 9057 | 397 | 4,38 |
| 21.4.2007 | 6761 | 217 | 3,21 |
| 22.4.2007 | 9 336 | 315 | 3,37 |
| 23.4.2007 | 9 428 | 292 | 3,10 |
| 24.4.2007 | 10 441 | 228 | 2,18 |
| 25.4.2007 | 7 530 | 284 | 3,77 |
| 26.4.2007 | 7 645 | 209 | 2,73 |
| 27.4.2007 | 7 884 | 265 | 3,36 |
| 28.4.2007 | 4 002 | 187 | 4,67 |
| 29.4.2007 | 4 653 | 182 | 3,91 |
| 30.4.2007 | 6 186 | 150 | 2,42 |
| 1.5.2007 | 6 325 | 218 | 3,45 |
| 2.5.2007 | 6 426 | 173 | 2,69 |
| 3.5.2007 | 7 910 | 172 | 2,17 |
| 4.5.2007 | 8 800 | 315 | 3,58 |

| Datum | DUNNO | REJECT | % |
|---|---|---|---|
| 5.5.2007 | 8 773 | 238 | 2,71 |
| 6.5.2007 | 8 663 | 256 | 2,96 |
| 7.5.2007 | 9 274 | 267 | 2,88 |
| 8.5.2007 | 6 786 | 201 | 2,96 |
| 9.5.2007 | 7 119 | 206 | 2,89 |
| 10.5.2007 | 7 794 | 202 | 2,59 |
| 11.5.2007 | 5 779 | 201 | 3,48 |
| 12.5.2007 | 5 272 | 209 | 3,96 |
| 13.5.2007 | 7 695 | 195 | 2,53 |
| 14.5.2007 | 10 972 | 651 | 5,93 |
| 15.5.2007 | 8 796 | 501 | 5,70 |
| 16.5.2007 | 7 244 | 229 | 3,16 |
| 17.5.2007 | 7 759 | 277 | 3,57 |
| 18.5.2007 | 10 724 | 467 | 4,35 |
| 19.5.2007 | 7 745 | 324 | 4,18 |
| 20.5.2007 | 6 299 | 245 | 3,89 |
| 21.5.2007 | 6 305 | 206 | 3,27 |

Tabulka 1: Vyhodnocení kontroly SPF

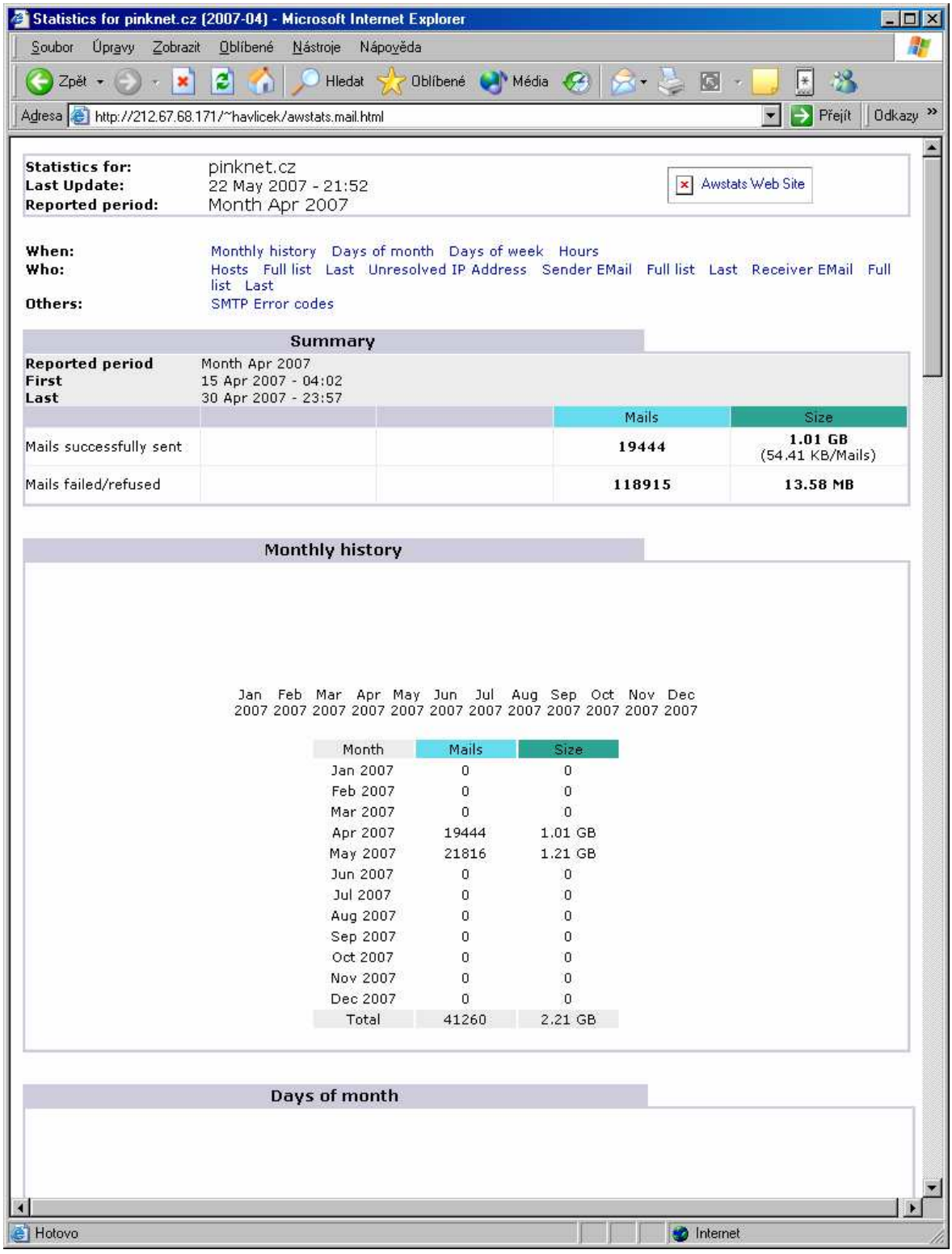
Z uvedené tabulky vidíme, že podíl počtu odmítnutých zpráv z počtu zpráv vyhodnocených s neutrálním výsledkem se pohybuje od přibližně 2 % až do téměř 6 %, v extrémním případě náhodného výkyvu dokonce 8 %.

Dle názoru autora je vzhledem k jednoduchosti a nenáročnosti na strojový čas u této kontroly výsledek uspokojivý a provádět testování na validitu SPF má smysl. Zároveň lze předpokládat, že se vzrůstajícím množstvím domén deklarujících SPF záznamy bude účinnost testování SPF vzrůstat.

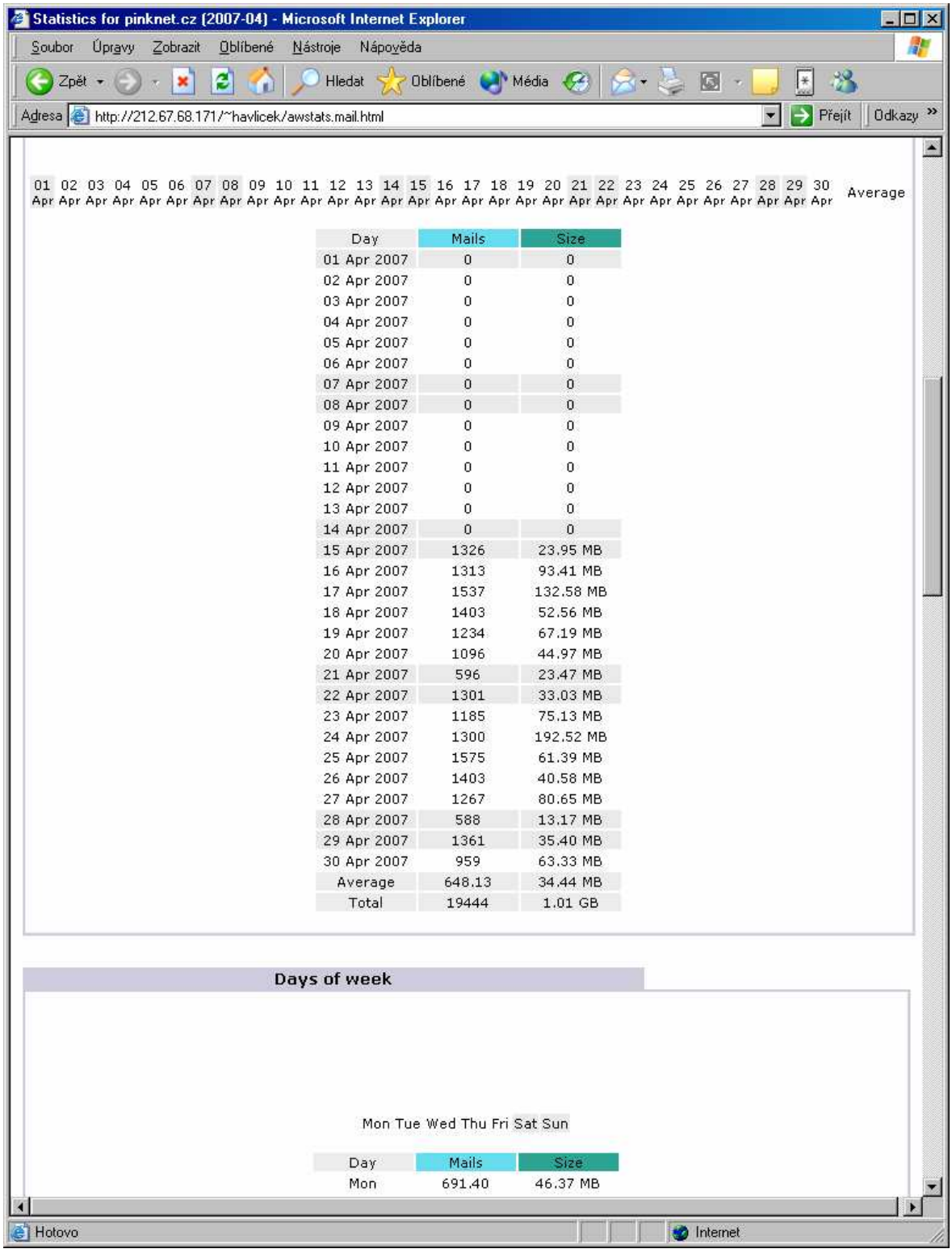
## 9.2. Vyhodnocení ochrany metodou greylistingu

Účinnost antispamové ochrany metodou greylistingu doložíme opět vyhodnocením emailového provozu domény pinknet.cz za období 15.4.2007 až 21.5.2007. Dále zhodnotíme graf statistiky záchytu antivirového programu provozovaného na centrálním mailserveru Vysoké školy ekonomické v Praze za rok 2004, kdy zde byl greylisting nasazen.

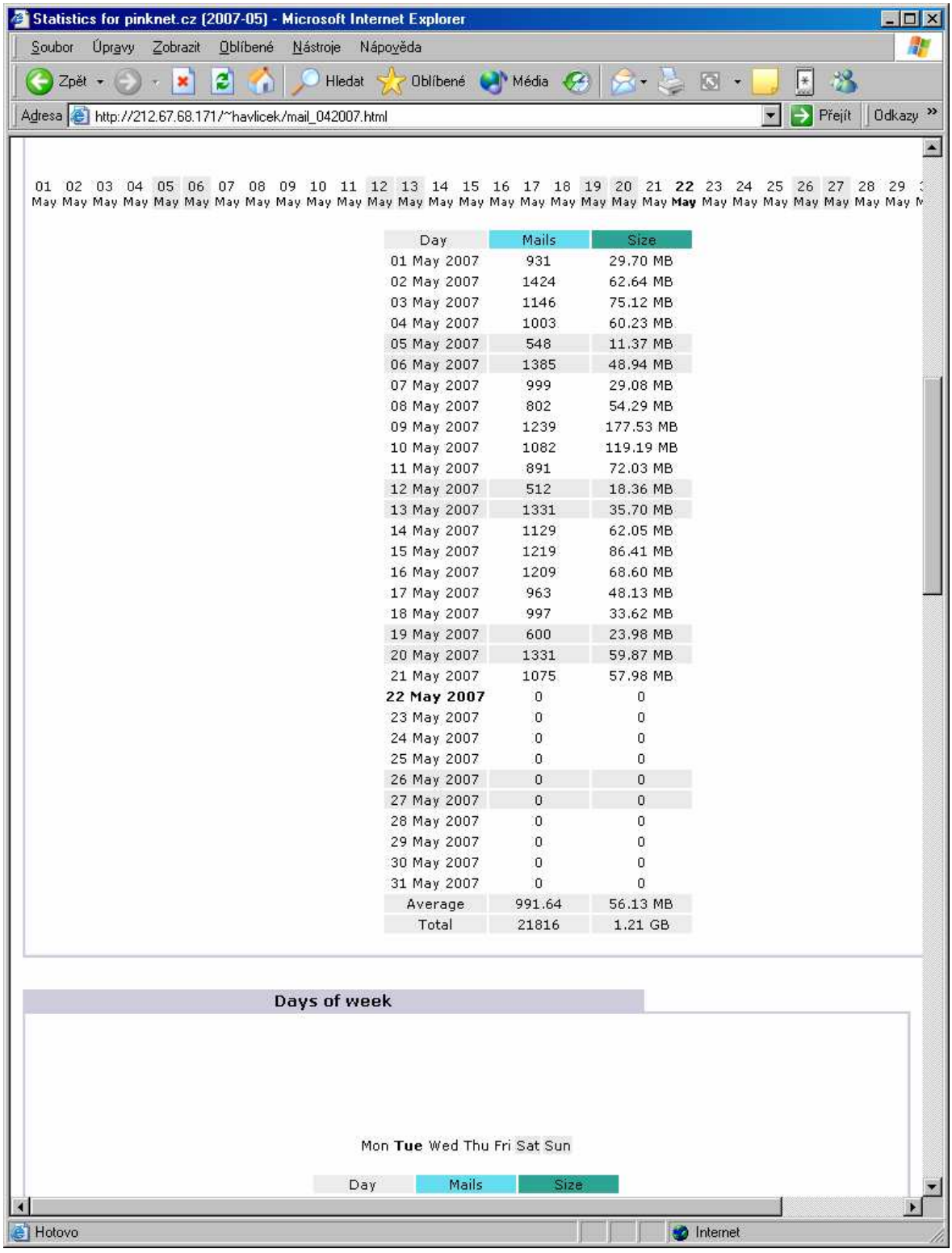
Logy emailového provozu domény pinknet.cz jsou pro získání souhrnných údajů zpracovány programem AWStats [19]. Pomocí tohoto SW získáme přehled o uskutečněném emailovém provozu po jednotlivých dnech, doménách odesílatelů a příjemců a další podrobnosti. Bohužel chybové návratové kódy mailserveru tento program reportuje pouze souhrnně, a nelze tedy stoprocentně určit úspěšnost greylistingu jako takového.




Obrázek 2: Celkové statistiky emailového provozu domény pinknet.cz v programu AWStats za duben a květen 2007




Obrázek 3: Statistiky emailového provozu domény pinknet.cz v programu AWStats – duben 2007




Obrázek 4: Statistiky emailového provozu domény pinknet.cz v programu AWStats – květen 2007

Číselné údaje celkem doručených zpráv, celkové velikosti a počty odmítnutí zprávy filtrem greylistingu jsou uvedeny v následující tabulce:

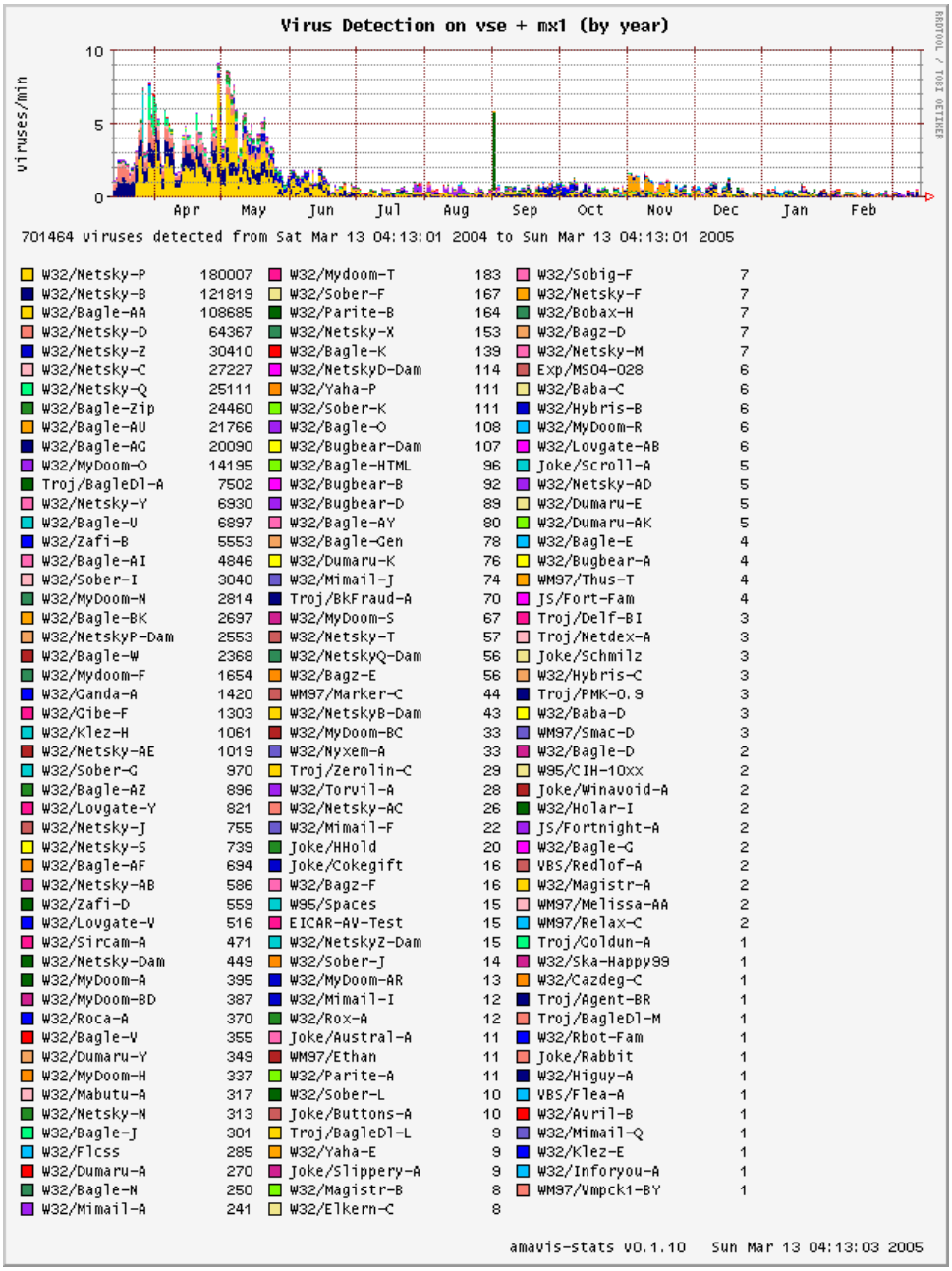
| Datum | Počet zpráv | Celková velikost | Počet odmítnutí greylistingem |
|---|---|---|---|
| 15 Apr 2007 | 1326 | 23.95 MB | 4153 |
| 16 Apr 2007 | 1313 | 93.41 MB | 8356 |
| 17 Apr 2007 | 1537 | 132.58 MB | 7698 |
| 18 Apr 2007 | 1403 | 52.56 MB | 7669 |
| 19 Apr 2007 | 1234 | 67.19 MB | 8195 |
| 20 Apr 2007 | 1096 | 44.97 MB | 7781 |
| 21 Apr 2007 | 596 | 23.47 MB | 6021 |
| 22 Apr 2007 | 1301 | 33.03 MB | 8423 |
| 23 Apr 2007 | 1185 | 75.13 MB | 7932 |
| 24 Apr 2007 | 1300 | 192.52 MB | 8815 |
| 25 Apr 2007 | 1575 | 61.39 MB | 5983 |
| 26 Apr 2007 | 1403 | 40.58 MB | 6154 |
| 27 Apr 2007 | 1267 | 80.65 MB | 6423 |
| 28 Apr 2007 | 588 | 13.17 MB | 3334 |
| 29 Apr 2007 | 1361 | 35.40 MB | 3979 |
| 30 Apr 2007 | 959 | 63.33 MB | 4879 |
| 01 May 2007 | 931 | 29.70 MB | 5334 |
| 02 May 2007 | 1424 | 62.64 MB | 4913 |
| 03 May 2007 | 1146 | 75.12 MB | 6476 |
| 04 May 2007 | 1003 | 60.23 MB | 7311 |
| 05 May 2007 | 548 | 11.37 MB | 7971 |
| 06 May 2007 | 1385 | 48.94 MB | 7835 |
| 07 May 2007 | 999 | 29.08 MB | 8125 |
| 08 May 2007 | 802 | 54.29 MB | 5758 |
| 09 May 2007 | 1239 | 177.53 MB | 5831 |
| 10 May 2007 | 1082 | 119.19 MB | 6526 |
| 11 May 2007 | 891 | 72.03 MB | 4727 |
| 12 May 2007 | 512 | 18.36 MB | 4585 |
| 13 May 2007 | 1331 | 35.70 MB | 6940 |
| 14 May 2007 | 1129 | 62.05 MB | 9337 |
| 15 May 2007 | 1219 | 86.41 MB | 7176 |
| 16 May 2007 | 1209 | 68.60 MB | 5736 |
| 17 May 2007 | 963 | 48.13 MB | 6447 |
| 18 May 2007 | 997 | 33.62 MB | 9442 |
| 19 May 2007 | 600 | 23.98 MB | 6966 |
| 20 May 2007 | 1331 | 59.87 MB | 5583 |
| 21 May 2007 | 1075 | 57.98 MB | 4850 |

Tabulka 2: Počty doručených zpráv a počty odmítnutí greylistingem

Údaje v tabulce 2 nelze nicméně vyhodnocovat procentuálním poměrem odmítnutých a doručených zpráv z několika důvodů. V první řadě je třeba vzít v úvahu, že program AWStats započítává do celkového počtu zpráv emaily doručené oběma směry, tedy přijaté i odeslané. Zprávy odeslané nicméně greylistingem kontrolovány nejsou, neboť to nemá žádný smysl. Dále je třeba zohlednit skutečnost, že z tabulky nijak neplyne, kolik pokusů o

doručení bylo po odmítnutí s úspěchem opakováno. Počty odmítnutých zpráv tedy nelze ztotožňovat s počty úspěšně odstraněných spamů. Není ani nijak zohledněn primitivní whitelisting na ukázkovém serveru používaný.

Dalším podkladem pro zhodnocení účinnosti greylistingu je graf zachycených virů na centrálních mailserverech Vysoké školy ekonomické v Praze, převzatý z Výroční zprávy Výpočetního centra VŠE v Praze za rok 2004 [20], uvedený na obr. 5. Na obrázku je patrný razantní pokles antivirem zachycených virů v období květen až červen 2004. Toto období koresponduje s testovacím nasazením greylistingu na VŠE v Praze na konci května, v průběhu června byl proces laděn a došlo ke změně a úpravám v té době použitého policy serveru. Z toho důvodu nebyl greylisting v provozu neustále. Od počátku července byl již greylisting nasazen stabilně.




Obrázek 5: Statistika záchytu virů antivirovým systémem na centrálních mailserverech VŠE v Praze v roce 2004

Účinnost greylistingu lze hodnotit empiricky na základě zkušeností. V článku Greylisting aneb kladivo na spam [21] Petr Krčmář uvádí svoji zkušenost s touto metodou boje proti spamu. Zpožďování pošty považuje za snesitelné, pouze mírně nepříjemné. Uvádí, že z cca 80 spamů došlých před nasazením greylistingu každý den jich po aplikaci metody zbude asi šest.

Zkušenost autora práce je následující. Na emailovou adresu v doméně vse.cz, existující od konce roku 1995 a používanou v různých diskusních skupinách a usenet news v době, kdy

spam nebyl problémem, docházelo autorovi na počátku roku 2004 cca 60 – 100 spamů denně. Část z nich byla zachycena v té době používaným spamassasinem a část zůstala neoznačena a musela být mazána ručně. Po několikadenní nepřítomnosti to obvykle znamenalo až několikahodinovou práci. Po nasazení greylistingu v roce 2004 počet doručených spamů na stejnou adresu poklesl na 0 – 3 denně s téměř stoprocentním záchytem spamassasinem. Spam tedy téměř není nutné řešit ručně. Tento stav od června roku 2004 trvá doposud (květen 2007), bez nutnosti jakkoliv upravovat použitý greylistovací algoritmus nebo jeho parametry. Spameři se tedy greylistingu do této doby nepřizpůsobili. Po nasazení greylistingu na VŠE dokonce docházelo k situaci, kdy si někteří uživatelé stěžovali, že jim přestala docházet elektronická pošta.

Z výše uvedeného vyplývá, že greylisting lze považovat za metodu s vysokou účinností a nízkým sklonem k odmítání korektních zpráv. Vzhledem k jisté kontroverznosti, spočívající ve zvýšené zátěži odesílající strany, je vhodné jej kombinovat s udržovaným whitelistem.

## 9.3. Zhodnocení nasazené ochrany z hlediska bezpečnosti

Z pohledu bezpečnostní politiky IT infrastruktury jsou s provozem emailu spojovány zejména dvě hrozby. V první řadě jde o možný kanál úniku citlivých informací. Nasazením ochrany spočívající ve striktním vyžadování autorizace při odesílání pošty a blokování portu 25 na hraničním routeru je zabráněno náhodnému nebo cíleně na dálku vyvolanému odeslání citlivých dat pomocí elektronické pošty. Z tohoto pohledu lze antispamovou ochranu považovat za přínosnou a z bezpečnostního hlediska pozitivní.

Druhou obvyklou hrozbou je riziko zavlečení škodlivého kódu buď přímo zasláným emailem (virus nebo worm útočící například na zranitelnost MUA), nebo infekcí prostřednictvím WWW stránek, na které adresáta spam navádí například za účelem phishingu. Nasazení účinných antispamových technik tedy tuto hrozbu významně omezuje, a jde tedy též o pozitivum.

Obecně je z pohledu bezpečnostní politiky třeba vždy mít na zřeteli, že elektronická pošta založená na protokolu SMTP je z principu nespolehlivá a nedůvěryhodná. Tyto vlastnosti nelze zvrátit žádnými antispamovými opatřeními, lze je pouze zmírnit používáním elektronických podpisů a šifrování zpráv. Protože ale podpisy a šifrování nelze striktně vyžadovat od okolního světa, riziko spamu lze takto pouze snížit.

# ZÁVĚR

Tato práce se zabývá v poslední době velmi rozšířeným fenoménem nevyžádaných zpráv v elektronické komunikaci, zejména v elektronické poště. V úvodu práce byl fenomén spamu definován a popsán. V teoretické části práce jsem popsal základní principy fungování elektronické pošty a zabýval se možnostmi ochrany proti spamu distribuovanému prostřednictvím elektronické pošty. Na dostupné metody lze pohlížet z různých úhlů. V práci se snažím problematiku antispamové ochrany nahlížet ze strany koncového uživatele v jednom pohledu a systémového administrátora, starajícího se o příjem i o odesílání elektronické pošty, v pohledu druhém. Vlastní techniky ochrany lze dělit na administrativně organizační a dále na technické, automatizované, založené na technické specifikaci komunikačního protokolu nebo na analýze vlastností a obsahu vlastní zprávy. U každé metody jsou diskutovány výhody i rizika.

Okrajově je teoretická část věnována i problematice spamu v jiných elektronických komunikačních kanálech, než je elektronická pošta.

V praktické části je popsán návrh řešení antispamové ochrany použitelný i pro rozsáhlou síť, ale svojí strukturou aplikovatelný i na síť komorních rozměrů. Je zde navrženo několik organizačních opatření, upravena architektura emailové infrastruktury a zvolen SW pro centrální emailový server.

Na zvolených prostředcích je následně navrženo použití antispamové ochrany pomocí metody kontroly SPF záznamů a v druhém kroku metody greylistingu. Je vybrán a doporučen vhodný doplňkový SW pro realizaci těchto opatření a obě metody jsou i zhodnoceny pomocí empiricky získaných dat z mailserveru, na jehož provozu se autor podílí, i z dalších zdrojů.

Domnívám se, že cílů práce definovaných v úvodu a zadání bylo úspěšně dosaženo.

# CONCLUSION

This diploma thesis deals with the widely spread phenomenon of unwanted messages in electronic communication, especially in the electronic mail. In the opening of the thesis, the phenomenon of spam is described and specified. The theoretical part of the thesis describes the basic principles of electronic mail operation and summarizes possibilities of protection against spam distributed through electronic mail.

Anti-spam methods can be viewed from different angles. The thesis looks at the anti-spam protection from the viewpoint of the end user, as well as from the viewpoint of the system administrator who takes care of the electronic mail traffic. The anti-spam techniques might be divided into two groups: administratively organizational techniques being the first group and the technical, automated techniques forming the second one. The latter mentioned techniques are based on the technical specification of the communication protocol and on the analysis of the actual message content and properties. The advantages as well as the risks of each method are discussed; moreover, the theoretical part touches upon the questions of spam in other channels of electronic communication as well.

The practical part of the thesis describes a model of anti-spam protection designed for a large network; however, its structure might be applied to a smaller network all the same. This part of the thesis contains a set of proposed organizational measures, adapted email infrastructure architecture and a description of the software chosen for the central mail server.

With selected resources, anti-spam protection through SPF record control method is proposed, the second step being the method of greylisting. Suitable additional software is selected and recommended for the implementation of these measures. Both methods are evaluated through empirically gathered data from the mailserver, the author of the thesis administers, as well as other resources.

The author of the thesis believes that the objectives stated in the assignment and in the introduction were successfully fulfilled.

# SEZNAM POUŽITÉ LITERATURY

[A] *Comer, Douglas,* Internetworking with TCP/IP Volume I: Principles, Protocols and Architecture 3rd. ed, Prentice Hall inc, 1885, 613 str., ISBN 0-13-216987-8

[B] *Roberts, Dave*, Internet protocols handbook, The Coriolis Group, 1996, 448 str., ISBN 1-883577-88-8

[C] *Rose, Marshall T.*, The INTERNET MESSAGE closing the book with electronic mail, PTR Prentice Hall, 1994, 370 str., ISBN 0-13-092941-7

[D] *Krol, Ed*, Vše o Internetu - Průvodce uživatele katalogových zdrojů, Science 1995 (překlad O'Reilly & Associates Inc. 1994], 490 str., ISBN 80-901475-4-2

[1] *Encyklopedie Wikipedia CS* [online]. 11.11.2006.[cit. 28.11.2006] Dostupný z WWW <http://cs.wikipedia.org/wiki/Spam>

[2] *Wiki Apache foundation* [online]. 10.7.2006. [cit. 29.11.2006] Dostupný WWW <http://wiki.apache.org/spamassassin/>

[3] *Wikipedia.org*, [online], 27.11.2006 [cit. 29.11.2006] Dostupný z WWW <http://en.wikipedia.org/wiki/Bayes'_theorem>

[4] *Kára, Michal,* Nepoužívejte IP blacklisty! (1.), Lupa.cz [online]. 20.5.2005 [cit. 4.12.2006], Dostupný z WWW,

<http://www.lupa.cz/clanky/nepouzivejte-ip-blacklisty-1/>, ISSN 1213-0702

[5] *Kára, Michal,* Nepoužívejte IP blacklisty! (2.), Lupa.cz [online]. 23.5.2005 [cit. 4.12.2006], Dostupný z WWW,
<http://www.lupa.cz/clanky/nepouzivejte-ip-blacklisty-2/>, ISSN 1213-0702

[6] *Satrapa, Pavel*, Greylisting: nová metoda boje proti spamu, Lupa.cz [online], 23.4.2004, [cit. 4.12.2006], Dostupný z WWW,
<http://www.lupa.cz/clanky/greylisting-nova-metoda-boje-proti-spamu/>,

ISSN 1213-0702

[7] *Bjarne Lundgren*, www.greylisting.org [online], 2004 [cit. 4.12.2006], Dostupný z WWW, <http://www.greylisting.org/>

[8] Habeas Inc, www.habeas.cz [online], [cit. 4.12.2006], Dostupný z WWW, <http://www.habeas.com/>, komerční presentace.

[9] *Wikipedia.org* [online].3.12.2006, [cit 4.12.2006], Dostupný z WWW, <http://en.wikipedia.org/wiki/Captcha>.

[10] *Chalupa, Pavel,* Konečné řešení v boji proti spamu v diskusních fórech, Root.cz [online], 4.12.2006 [cit. 4.12.2006], Dostupný z WWW, < http://www.root.cz/clanky/konecne-reseni-v-boji-proti-spamu-v-diskusnich-forech/>, ISSN 1212-8309

[11] *Postel, Jonathan B.*, RFC 821 SIMPLE MAIL TRANSFER PROTOCOL, Internet Engineering Task Force (IETF) [online], August 1982 [cit. 17.4.2007], Dostupný z WWW <http://tools.ietf.org/html/rfc821.html >

[12] *Klensin, J.*, RFC 2821 Simple Mail Transfer Protocol, Internet Engineering Task Force (IETF) [online], April 2001 [cit. 17.4.2007], Dostupný z WWW <http://tools.ietf.org/html/rfc2821.html >

[13] *Myers, J.*, RFC 2554 SMTP Service Extension for Authentication, Internet Engineering Task Force (IETF) [online], March 1999 [cit. 18.4.2007], Dostupný z WWW <http://tools.ietf.org/html/rfc2554>

[14] *Wikipedia.org* [online]. 28.4.2007 [cit. 4.5.2007], Dostupný z WWW <http://en.wikipedia.org/wiki/DNSBL>

[15] *Wikipedia.org* [online]. 28.4.2007 [cit. 4.5.2007], Dostupný z WWW <http://en.wikipedia.org/wiki/Comparison_of_DNS_blacklists>

[16] *Venema, Wietse*, The Postfix Home Page, [online], [cit 16.5.2007], Dostupný z WWW <http://www.postfix.org/>

[17] Sender Policy Framework, [online], 17.5.2007, [cit 21.5.2007], Dostupný z WWW <http://www.openspf.org/>

[18] *Bouton, Lionel*, SQLgrey Sourceforge project page, [online], [cit 21.5.2007], Dostupný z WWW <http://sqlgrey.sourceforge.net/>

[19] *Destailleur, Laurent,* AWStats official web page, [online], [cit 22.5.2007], Dostupný z WWW http://awstats.sourceforge.net/

[20] *Čermák, Igor, RNDr. za kolektiv*, Výroční zpráva za rok 2004 – Výpočetní centrum, Vysoká škola ekonomická v Praze - Nakladatelství Oeconomica, 2005, 125 str, ISBN 80-245-0922-9

[21] *Krčmář, Petr*, Greylisting aneb kladivo na spam, Root.cz, [online], 3.8.2006, [cit 22.5.2007],
Dostupný z WWW <http://www.root.cz/clanky/greylisting-aneb-kladivo-na-spam/> ISSN 1212-8309

Seznam použitých symbolů a zkratek

| | |
|---|---|
| CNAME | Common Name - "přezdívka" v DNS systému, symbolické jméno odkazující na jiné symbolické jméno |
| DNS | Domain Name System - systém překladu doménových jmen na IP adresy a zpět |
| HTTP | HyperText Transport Protocol - protokol určený k přenosu WWW stránek |
| IM | Instant Messaging - způsob online komunikace |
| IP | Internet protocol – protokol používaný na síťové vrstvě OSI modelu |
| ISP | Internet Service Provider - poskytovatel služeb Internetu, zejména připojení |
| MTA | Message Transport Agent - serverový proces pro předávání elektronické pošty |
| MUA | Message User Agent - klientský program pro práci s elektronickou poštou |
| MX | Mail Exchanger - záznam v DNS určující, jak má být doručována elektronická pošta |
| OCR | Optical Character Recognition - optické rozpoznávání znaků v obrázku, například po naskenování textu |
| PTR | Pointer Record - reverzní záznam v DNS překládající IP adresu na doménové jméno |
| RFC | Request For Comment - systém "norem" komunikace v Internetu |
| SEO | Search Engine Optimalization - optimalizace kódu WWW stránek tak, aby se s nimi lépe pracovalo vyhledávacím strojům, marketingový nástroj |
| SMTP | Simple Message Transport Protocol - protokol používaný pro přenos elektronické pošty |
| SPF | Sender Policy Framework - antispam technika |
| SSL | Secure Socket Layer - metoda šifrování dat pro přenos dat |
| TCO | Total Cost of Ownership – celkové náklady na vlastnictví technologie, zahrnují nejen pořizovací cenu, ale i cenu údržby, podpory a práce |

| | |
|---|---|
| URI | Uniform Ressource Identifier – metoda, jak jednoznačně identifikovat zdroje. |
| URL | Unifirm Ressource Locator – unifikovaná metoda jednoznačné lokalizace zdroje na Internetu |
| UUNET | Unix to Unix copy Network – dříve používaná asynchronní metoda předáváníd at |
| VoIP | Voice over IP - přenos hlasu prostřednictvím protokolu TCP/IP |
| VPN | Virtual Private Network – virtuální šifrované spojení mezi dvěma body sítě zajišťující bezpečný přenos dat |

# SEZNAM OBRÁZKŮ

# SEZNAM TABULEK

# SEZNAM PŘÍLOH

## Příloha P I: Struktura databáze policy serveru SQLgrey

# PŘÍLOHA P I: STRUKTURA DATABÁZE POLICY SERVERU SQLGREY

```
mysql> use sqlgrey;
Database changed
mysql> show tables;
+-------------------+
| Tables_in_sqlgrey |
+-------------------+
| config            |
| connect           |
| domain_awl        |
| from_awl          |
| optin_domain      |
| optin_email       |
| optout_domain     |
| optout_email      |
+-------------------+
8 rows in set (0.00 sec)

mysql> describe config;
+-----------+--------------+------+-----+---------+-------+
| Field     | Type         | Null | Key | Default | Extra |
+-----------+--------------+------+-----+---------+-------+
| parameter | varchar(255) | NO   | PRI |         |       |
| value     | varchar(255) | YES  |     | NULL    |       |
+-----------+--------------+------+-----+---------+-------+
2 rows in set (0.00 sec)

mysql> describe connect;
+---------------+--------------+------+-----+-------------------+-------+
| Field         | Type         | Null | Key | Default           | Extra |
+---------------+--------------+------+-----+-------------------+-------+
| sender_name   | varchar(64)  | NO   |     |                   |       |
| sender_domain | varchar(255) | NO   |     |                   |       |
| src           | varchar(39)  | NO   | MUL |                   |       |
| rcpt          | varchar(255) | NO   |     |                   |       |
| first_seen    | timestamp    | YES  | MUL | CURRENT_TIMESTAMP |       |
+---------------+--------------+------+-----+-------------------+-------+
5 rows in set (0.01 sec)

mysql> describe domain_awl;
+---------------+--------------+------+-----+---------------------+-------+
| Field         | Type         | Null | Key | Default             | Extra |
+---------------+--------------+------+-----+---------------------+-------+
| sender_domain | varchar(255) | NO   | PRI |                     |       |
| src           | varchar(39)  | NO   | PRI |                     |       |
| first_seen    | timestamp    | YES  |     | CURRENT_TIMESTAMP   |       |
| last_seen     | timestamp    | YES  | MUL | 0000-00-00 00:00:00 |       |
+---------------+--------------+------+-----+---------------------+-------+
4 rows in set (0.00 sec)

mysql> describe from_awl;
+---------------+--------------+------+-----+---------------------+-------+
| Field         | Type         | Null | Key | Default             | Extra |
+---------------+--------------+------+-----+---------------------+-------+
| sender_name   | varchar(64)  | NO   | PRI |                     |       |
| sender_domain | varchar(255) | NO   | PRI |                     |       |
| src           | varchar(39)  | NO   | PRI |                     |       |
| first_seen    | timestamp    | YES  |     | CURRENT_TIMESTAMP   |       |
| last_seen     | timestamp    | YES  | MUL | 0000-00-00 00:00:00 |       |
+---------------+--------------+------+-----+---------------------+-------+
5 rows in set (0.00 sec)
```

```
mysql> describe optin_domain;
+--------+--------------+------+-----+---------+-------+
| Field  | Type         | Null | Key | Default | Extra |
+--------+--------------+------+-----+---------+-------+
| domain | varchar(255) | NO   | PRI |         |       |
+--------+--------------+------+-----+---------+-------+
1 row in set (0.00 sec)

mysql> describe optin_email;
+-------+--------------+------+-----+---------+-------+
| Field | Type         | Null | Key | Default | Extra |
+-------+--------------+------+-----+---------+-------+
| email | varchar(255) | NO   | PRI |         |       |
+-------+--------------+------+-----+---------+-------+
1 row in set (0.00 sec)

mysql> describe optout_domain;
+--------+--------------+------+-----+---------+-------+
| Field  | Type         | Null | Key | Default | Extra |
+--------+--------------+------+-----+---------+-------+
| domain | varchar(255) | NO   | PRI |         |       |
+--------+--------------+------+-----+---------+-------+
1 row in set (0.00 sec)

mysql> describe optout_email;
+-------+--------------+------+-----+---------+-------+
| Field | Type         | Null | Key | Default | Extra |
+-------+--------------+------+-----+---------+-------+
| email | varchar(255) | NO   | PRI |         |       |
+-------+--------------+------+-----+---------+-------+
1 row in set (0.00 sec)
```