

## 7.2 Unterdrückung von Spam E-Mails

Vor einiger Zeit hätte an dieser Stelle noch eine umfangreiche Erklärung gestanden, was Spam E-Mails sind und woher dieses Wort überhaupt stammt. Angesichts der realen Verhältnisse Anno 2003 mit einem Anteil von Spam E-Mails von 30% bis 50% braucht niemand mehr eine entsprechende Belehrung. Spam ist zum Alltag geworden; die Bekämpfung bzw. Unterdrückung von Spam E-Mails zu einem Bedürfnis.

Dabei möchte ich nicht die Berechtigung des kommerziellen Einsatzes von E-Mails in Frage stellen und diese von Spam E-Mails abgrenzen. Um kommerzielle E-Mails zu erhalten, bedienen sich die Anbieter im wesentlichen folgender Methoden:

Kommerzielle E-Mails

- **Opt-In** — Der Interessent muss sich z.B. über ein Webformular mit seiner E-Mail-Adresse in eine E-Mail-Verteilerliste einschreiben. Der Anbieter verschickt üblicherweise anschliessend eine Bestätigungs-E-Mail an die eingetragene E-Mail-Adresse. Nachteil diese Verfahrens ist, dass natürlich auch Dritte unter falschen Namen und E-Mail-Adresse sich eintragen können (Adressen-Spoofing).
- Daher bieten alle Anbieter auch ein **Opt-Out** an, mit man sich von der Liste wieder löschen kann. Häufig wird auch bei Spam E-Mails eine Opt-Out-Möglichkeit angeboten; in aller Regel ist dies nicht seriös gemeint sollte daher nie genutzt werden (siehe unten).
- **Double Opt-In** — Hierbei trägt sich der Interessent zunächst wie beim Opt-In in eine E-Mail-Verteilerliste ein; diese verschickt aber dann nicht eine einfache Bestätigungs-E-Mail sondern ein sog. Cookie, dass der Absender bestätigen muss. Erst in diesem Falle wird er in die Verteilerliste aufgenommen. Vorteil hierbei ist, dass sowohl die Gültigkeit der E-Mail-Adresse des Interessenten überprüft, als auch ein mögliches Spoofing durch Dritte unterbunden wird.

### 7.2.1 Unsolicited Commercial E-Mails (UCE)

In Abgrenzung hierzu, haben wir es bei Spam um sog. Unsolicited Commercial E-Mails zu tun. Tatsache ist jedenfalls, dass der Umfang in den letzten Jahren/Monaten im allgemeinen stark zugenommen hat. Die individuelle Belästigung durch Spam schwankt jedoch beträchtlich; und zwar sowohl hinsichtlich der emotionalen Betroffenheit als auch hinsichtlich der Anzahl der Spam E-Mails. Es gibt zwar einige Aussagen hinsichtlich des Umfangs von Spam bei der Internet E-Mail, aber die Schwankungsbreite dieser Zahlen ist beträchtlich und wird heute zwischen 30% und 60% abgeschätzt, wobei in der Regel nicht definiert wurde, wie diese Zahlen ermittelt wurden und wie sie zu interpretieren ist.

### 7.2.1.1 Ursachen

Spam E-Mails dienen in der Regel (> 99%) zur Anbiederung kommerzieller Dienste oder Produkte. Ganz vorne in der Statistik agieren Porno E-Mails (direkt oder indirekte "scharfe Fotos auf meiner Web-Site"), gefolgt von diversen pharmazeutisch-medizinischen Hilfsmitteln à la Viagra, Penis- und Brustvergrößerungen, der bekannten "Blue Pill" (nach einem Song der Bangels), in Deutschland 0190-Dialern, Anti-Spam Tools (!) und natürlich der "Nigeria-Connection".

Gemeinsam ist, dass trotz minimalster Trefferraten und Effizienz, die Gewinnmargen (besser gesagt: das Leimen der Betroffenen) so lukrativ ist, dass sich solche Massen-E-Mails lohnen. Die Trefferraten bewegen sich häufig im Pörmille-Bereich; wieviel der E-Mail-Anwender dann hierauf reagieren, ist unklar. In Anbetracht des wachsenden Verdrusses über Spam E-Mails sinkt dieser Anteil sicherlich (ein Alleinstellungsmerkmal fehlt), dem die Spammer durch höhere Trefferraten begegnen wollen ("Brut-Force-Attacks"), was das ganze System hochschaukelt.

### 7.2.1.2 Herkunft

Soweit bekannt ist, stammen die meisten Spam E-Mails aus USA, Brasilien, Korea und Deutschland. Verfolgen kann man den Weg der Spam E-Mails an der Reihenfolge der "Received:" Zeilen im Header der E-Mail. Hier ist besonders der erste (d.h. unterste) Eintrag von Belang, da dies die erste Station der Spam E-Mail darstellt. Ob allerdings in der Zukunft die Spammer nicht dazu übergehen, diese zu fälschen (siehe weiter unten), ist allerdings unklar.

Spammer nehmen die Erzeugung und den Versand ihrer Spam E-Mails in wachsendem Masse professionell vor:

- Für den Zweck des Spam-Versands werden eigene Rechner aufgebaut (und deren "Baurezept" teilweise in den Spam E-Mails selbst offen beworben wird), um sie nach Ende der Spam-Aktion wieder abzubauen.
- Die Spam-Absender nutzen schnelle Internet-Anbindungen (z.B. DSL) über die eine grosse Anzahl von Spam E-Mails in kurzer Zeit versandt werden können. Wir reden hier von einem erzeugten E-Mail Aufkommen von Millionen/Stunde.
- Alternativ werden offene E-Mail-Server ("Open Relays") sowie offene Proxis genutzt, die vorher gescannt, als solche identifiziert und auf ihre Relaying-Eigenschaft getestet wurden.

Die Herkunft der E-Mail — kenntlich durch die SMTP "Mail From:" Adresse — ist in aller Regel gefälscht (*forged*), um möglichst unverdächtig zu erscheinen. Häufig werden Adressen der grossen E-Mail Internet-Anbieter wie "yahoo.com",

"yahoo.de", "web.de", "hotmail.com" und "gmx.de" gefälscht, wobei dies davon abhängig gemacht wird, in welches Land (per Domain-Kennung) die Spam E-Mail versandt wird. Der Benutzername, d.h. der Teil der E-Mail-Adresse links vor dem "@", wird meist über eine Zufallsfolge von Buchstaben von Zahlen generiert; was durchaus auch bei normalen Anwendern nicht unüblich ist.

### 7.2.1.3 Adressen und Adresshandel

Den Umfang der Spam E-Mails sowohl nach Anzahl wie nach übertragenem Volumen abzuschätzen, ist in aller Allgemeinheit nicht möglich. An dieser Stelle wage ich zu unterscheiden zwischen:

- Öffentlichen Institutionen, die einen Teil ihrer E-Mail-Adressen veröffentlichen; sei es über das Web, über Newsgroups, LDAP, X.500 oder über andere Medien.
- Firmen, deren E-Mail-Adressen in der Regel nur Kunden bekannt sind und die darüber hinaus "funktionale" Adressen öffentlich machen (Postmaster, Info, Vertrieb, Hotline etc.).
- Anwender mit privaten E-Mail-Adressen bei unterschiedlichen ISPs und E-Mail-Providern.
- Anwender, mit E-Mail-Adressen bei den sog. Free-Mailern wie z.B. GMX, Web.de, Yahoo.de oder anderen.
- Anwender, die sich an öffentlichen Foren im Internet beteiligen und dort registriert sind.

Um überhaupt an brauchbare E-Mail-Adressen zu kommen, bedienen sich die Spammer mehrere Methoden (vgl. <http://www.rehbein-dortmund.de/spamtrap.html>):

- **Adresshandel** — Der derzeitige Preis für eine E-Mail-Adresse beträgt etwa 0,002 Cent; also 5 Millionen Adressen für 99 \$. Entsprechende Datensätze werden selbst als Spam beworben; teilweise mit entsprechender E-Mail Software dazu, um schnell Spam E-Mails versenden zu können. Inwieweit die kommerziellen Internet-Händler (wie Amazon) ihre Adressen an Dritte vermarkten, bleibt im Verborgenen. Einerseits garantieren diese natürlich einen Kundenschutz (bereits in ihrem eigenen Interesse); andererseits sind speziell ihre Adressen (mit den entsprechenden Kundenprofile) von grossem Interesse und daher sehr "wertvoll".
- **Harvesting** — Internet Web-Seiten sowie Internet-Foren wie Newsgroups werden über sog. E-Mail Spider auf verwertbare Adressen untersucht (*Data Mining*) und diese gesammelt.
- **Fake FTP** — Bei der Verbindung auf eine Web-Seite wird versucht, den

eingebauten FTP-Client des Web-Browsers zu starten; in der Regel antwortet dieser darauf mit ein sog. Anonymous-Login mit der E-Mail-Adresse des Anwenders als Passwort.

- **Rücklauf** — "Gold Plated" Adressen erhält der Spammer, falls tatsächlich einer der Adressaten von der "Opt-Out" Möglichkeit Gebrauch macht.

Neben der Werbung per Spam E-Mail ist also der Adresshandel im Internet zu einem lukrativen Geschäft geworden. Dies gilt im Besonderen in dem Kontext, dass speziell die E-Mail-Adressen der grossen Freemail-Provider schnell an Aktualität verlieren. Das ständige Katz-und-Maus Spiel erzeugt somit seine eigenen Regeln.

## 7.2.2 Arbeitsweise der Spammer

### 7.2.2.1 Temporäre Domains

Der einfachste und sicherste Weg für einen potentiellen Spammer ist es, eine oder gleich mehrere Domänen bei einem NIC anzumelden, einen Rechner mit E-Mail-Software aufzusetzen und über ein Skript die eingespeicherten E-Mail-Adressen mit dem vorgegeben Text abzusetzen.

Innerhalb einer Stunde können so mehrere Millionen E-Mails abgesetzt werden. Nach dem Ende der Aktion werden Domains abgemeldet oder einfach "leer" gelassen und der Rechner von der (DSL-) Leitung abgeklemmt und abgeschaltet ("Throwaway Domain").

Laut einer E-Mail von Markus Stumpf (SpaceNet) wurden Anfang Juli 2003 u.a. folgende temporäre Domains zum Absenden von Spam E-Mails benutzt:

```
" @brightbernies.us
  @brightdons.us
  @brighthenrys.us
  @brightsandys.us
  @dazzlingdans.us
  @fabulousozzies.us
  @famousmickeys.us
  @fantasticteds.us
  @friendlyhenrys.us
  @friendlyozzies.us
  @friendlysandys.us
  @friendlyvinnies.us
  @heathersholly.us
  @mightydans.us
  @robsrules.us
  @smartdons.us
  @smarthenrys.us
```

Beispiel für temporäre  
Domains

```
@wittyozzies.us
@wranglerron.us

[ ... ] this list is by far incomplete
```

We got spam from these domains today and the last few days.

The spammers obviously registered these domain, they have DNS servers for them and even MX and A records. The TTL of these records is one hour. All of these are registered by

```
Domain Name:          FRIENDLYSANDYS.US
Domain ID:            D4426467-US
Sponsoring Registrar: GO DADDY SOFTWARE, INC.
Domain Status:       ok
Registrant ID:       GODA-03439749
Registrant Name:     steven little
Registrant Organization: Unknown
Registrant Address1: 321 s college
Registrant City:     seattle
Registrant State/Province: Washington
Registrant Postal Code: 98144
Registrant Country:  United States
Registrant Country Code: US
Registrant Phone Number: +1.2063503057
Registrant Email:    stevenlittle206@yahoo.com
```

[ ... ]

```
Name Server:         NS1.ENCHANTINGIDEAS.NET
Name Server:         NS2.ENCHANTINGIDEAS.NET
Created by Registrar: GO DADDY SOFTWARE, INC.
Last Updated by Registrar: GO DADDY SOFTWARE, INC.
Domain Registration Date: Tue Jul 01 04:19:58 GMT 2003
Domain Expiration Date:  Wed Jun 30 23:59:59 GMT 2004
Domain Last Updated Date: Tue Jul 01 04:29:52 GMT 2003
```

Get some millions throwaway accounts for as cheap as 5 USD."

### 7.2.2.2 Open Relays

Um eigene Ressourcen zu schonen und die Effizienz zu erhöhen, benutzen Spammer sog. offene Relays — also SMTP-Server, die ein unbeschränktes Weitersenden von E-Mails an beliebige Empfänger zulassen. Qmail, z.B. agiert als offenes Relay wenn weder die Datei `./rcpthosts` noch `./morercpthosts`

vorhanden bzw. befüllt ist.

Jeder verantwortliche System-Administrator sichert natürlich sofort seinen E-Mail-Server gegen unbeabsichtigtes Relaying ab. Mit der Verbreiterung des Internets bis in den letzten Winkel der Welt, werden jedoch insbesondere Web- und E-Mail-Server aufgesetzt, wo dieses Paradigma nicht gilt. "Berüchtigt" in diesem Zusammenhang sind z.B. entsprechende Server in Korea, die in Schulen betrieben werden. Verfügt die Schule dann auch noch über einen schnellen Internet-Zugang, eignen sich diese Server als ausgezeichnete Relay-Systeme Spam E-Mails zu verschicken. Da für den mittelbar Betroffenen auch kein Schaden entsteht (ausser der Nutzung seiner Bandbreite) bleibt das Relaying auch folgenlos — entsprechend gering fällt die Verantwortlichkeit gegenüber der "Internet-Community" bzw. den unmittelbar Betroffenen aus.

### 7.2.2.3 Open Proxies

Im Zeitalter der DSL-Flat-Rate und der Home-LANs bauen sich viele Windows-Anwender permanent arbeitende SMTP-Proxies auf, über die E-Mails empfangen und versendet werden. PCs unter Windows XP Home Edition besitzen prinzipiell die gleichen TCP/IP-Fähigkeiten wie Unix-Workstations — auch hinsichtlich der Leistung.

Falsch oder ungenügend konfigurierte SMTP-Proxies lassen sich über einen IP/Port-Scanner ausfindig und für Spammer nutzbar machen. Besonders die Windows-Software "AnalogX" hat sich hier unrühmlich hervorgetan. Unter Unix stellen schlecht-konfigurierte Apache Web-Server ein grosses Risiko dar. Mit wenigen Ausnahmen besteht kein Grund darin, einen Web-Server als Proxy zu betreiben. In der Konfigurationsdatei `httpd.conf` ist (wie im Beispiel) folgende Passage auszukommentieren:

```
#
# Proxy Server directives. Uncomment the following lines to
# enable the proxy server:
#
#<IfModule mod_proxy.c>
#   ProxyRequests On
#   <Directory proxy:*>
#       Order deny,allow
#       Deny from all
#       Allow from .your-domain.com
#   </Directory>
#
# Enable/disable the handling of HTTP/1.1 "Via:" headers.
# ("Full" adds the server version; "Block" removes all
outgoing Via: headers)
# Set to one of: Off | On | Full | Block
```

```
#
# ProxyVia On
#
# To enable the cache as well, edit and uncomment the
following lines:
# (no cacheing without CacheRoot)
#
# CacheRoot "/usr/local/www/proxy"
# CacheSize 5
# CacheGcInterval 4
# CacheMaxExpire 24
# CacheLastModifiedFactor 0.1
# CacheDefaultExpire 1
# NoCache a-domain.com another-domain.edu joes.garage-
sale.com
#</IfModule>
# End of proxy directives.
```

Ist einmal festgestellt, dass für eine IP-Adresse das Port 25 (SMTP) offen ist, kann sehr einfach getestet werden, ob über diesen Proxy E-Mails relaying möglich ist.

#### 7.2.2.4 Trojaner, Malicious Code und Backdoors

Bei einigen Web-Servern (z.B. Apache 1.3.12) es möglich, über einen POST-Befehl und einem Buffer-Overflow (z.B. im PHP), Programme ablaufen zu lassen. Hierdurch können auch Spam E-Mails verschickt werden. Dazu kann entweder das Standard-Socket-Interface oder aber — über die Shell — ein Standard-Mail-Client wie **mail** oder **mailx** eingesetzt werden.

Gleiches ermöglichen Trojaner unter Windows, die sich als unverdächtige Programme tarnen und über das Web heruntergeladen werden können (z.B. als Dailer oder Programm zur angeblichen Performance-Optimierung). Die eigentliche Implementierung eines SMTP-Clients bedarf nur weniger Zeilen (Malicious) Code und reduziert sich auch hier auf einen geeigneten Socket-Aufruf.

Besonders problematisch ist hierbei, dass nun die E-Mails aus dem eigentlich geschützten Intranet-Bereich abgesetzt werden. Dies kann dazu führen kann, dass die Mail-Server der Firma auf einer Open-Relay Liste im Internet auftauchen, obwohl sie nicht die eigentlichen Verursacher sind, sondern nur Vermittler.

## 7.2.3 Merkmale von Spam E-Mails

### 7.2.3.1 Aufbau von Spam E-Mails

Das eigentliche Merkmal von Spam E-Mails ist, dass sie versucht keine identifizierbaren Merkmale aufzuweisen. SMTP-Absender, SMTP-Sender, Mail-Sender und die Angabe des "Subjects:" werden in einer Spam-Kampagne häufig gewechselt.

Bei mir flatterte im Juli 2003 folgende Spam E-Mail in den Postkasten:

**A few moments of your time could save you thousands over the life of your loan! What are you waiting for?**

[Visit Now](http://mtggreat1.com/4/index.asp?RefID=588897)

pN2p

Hierbei dient "pN2p" als Serialisierungs-Information. Hinter der Angabe "Visit Now" versteckt sich die URL <http://mtggreat1.com/4/index.asp?RefID=588897>.

Der in X-HTML formulierte Text wurde mit einem erheblichen "Rauschen" aufbereitet, sodass es Systeme schwer haben, das Verhältnis zwischen Nutz- und Spam-Information zu bestimmen:

```
<x-html><HTML>
<!--16--><table align="center" bgcolor="ff2222" width="615"
border="2" cellspacing="12" cellpadding="5"><tr>
<td bgcolor="ffffff" align="center"><font size="4"
face="Arial, Helvetica, sans-serif"><b>A few m<!--AQ-->oments
of your ti<!--tI-->me cou<!--cA-->ld sav<!--Fz-->e y<!--rT--
>ou t<!--NK-->housa<!--DX-->nds o<!--QJ-->ver the l<!--w2--
>ife of your loa<!--dq-->n! W<!--Ju-->hat are y<!--v9-->ou
wai<!--88-->ting for? <!--Le--><br><br><!--Ep--><A
href="http://mtggreat1.com/4/index.asp?RefID=588897"> <font
size="5">Vis<!--Bw-->it N<!--U7-->ow</font></a></font><!--Xs--
> </td><!--54--></tr><br></table><br><br><!--pR--> <br> <HTML>
```

Abschickt wird die Spam E-Mail mit folgenden Merkmalen:

| SMTP-Absender         | SMTP-Sender (IP) | Mail-Sender          | Subject:  |
|-----------------------|------------------|----------------------|-----------|
| elsey@tobe-online.com | 61.78.232.181    | kelsey <kelsey@tobe- | Congrats! |

|                      |                |                                    |                    |
|----------------------|----------------|------------------------------------|--------------------|
|                      |                | online.com>                        |                    |
| jeanine@crisp-mcs.ne | 200.83.243.55  | jeanine<br><jeanine@crisp-mcs.net> | Wanna know why?    |
| jeanine@crisp-mcs.ne | 61.255.105.123 | jeanine<br><jeanine@crisp-mcs.net> | Don't pass this up |

*Tabelle 7.1-1: Einige Merkmale einer typischen Spam E-Mail.*

Diese "harmlose" Spam E-Mail soll als Lehrbeispiel dienen, dass die Spam-Erzeuger/Sender es mittlerweile gelernt haben:

- Spam E-Mails von real existierenden (temporären) Domänen mit korrekter DNS-Infrastruktur zu versenden,
- die SMTP Envelope-Information gängigen RFCs zu entsprechen,
- den Aufbau des E-Mail-Headers strukturell untadelig zu gestalten,
- die Angabe des "Subjects:" möglichst unverdächtig abzufassen,
- mittels einer in den E-Mail Body eingebauten Serialisierungsinformation eine mögliche MD5-Checksum ins Leere laufen zu lassen und
- den Inhalt der E-Mail so zu optimieren, dass gängige Spam-Klassifizierungsmethoden (SpamAssassin) entweder getäuscht werden oder aber einen geringen "Score" ergeben.
- *Stealthing*: Hierunter verstehe ich die Visualisierung der eigentlichen Mitteilung durch Graphik oder ein anderes Objekt. Wird statt eines Text eine Graphik (mit gleichem textlichen Inhalt) eingebettet, ist dieser für die üblichen Spam-Klassifizierungsprogramme nicht auswertbar und die E-Mail kann nicht als Spam identifiziert werden.

Typischerweise würde trotzdem die E-Mail als Spam klassifiziert werden, da in ihr das Wort "loan" vorkommt — allerdings erst nachdem die HTML-Information entfernt wurde. Die Untersuchung des Textes mittels regulärer Ausdrücke auf den Begriff "loan" würde jedenfalls ins Leere laufen.

Es ist m.E. nur eine Frage der Zeit, bis sich die Spam-Absender an aktuelle (inhaltliche) Filterkriterien angepasst haben. Letztlich kann jede Spam E-Mail mittels eines Durchlaufs an die gerade gültigen SpamAssassin-Version so "getunt" werden, dass sie eine geringe Spam-Rating erzeugt: Hase-und-Igel-Spiel.

### 7.2.3.2 Fake Opt-In/Opt-Out

Ein Grossteil der Spam E-Mails lockt mit der Möglichkeit eines Opt-Out, d.h. Austragens aus der E-Mail Verteilerliste. Wir betrachten einmal einen Ausschnitt (in X-HTML formatiert) aus einer Spam E-Mail, die die Möglichkeit eines Opt-Out anbietet:

```
You are receiving this message as a member of the Opt-In  
America List. To remove your email address please  
click here  
We honor all remove requests.
```

Die "Opt-Out" Möglichkeit lautet in Klartext:

```
<H4>You are receiving this message as a member of the Opt-In<BR>  
America List. To remove your email address please<BR>  
<A HREF="http://www.kgbfvkk29y.com/www.youngforever22.com">  
click here</A><BR>We honor all remove requests.</H4>
```

Die Spam E-Mail gibt also vor, dass der Empfänger sich per Opt-In in der "America List" eingetragen hat und bietet ein Opt-Out. Die Webseite "www.youngforever22.com" ist tatsächlich aktiv und bietet dort eine "Remove" Möglichkeit, hinter der ein PHP-Skript steckt. Wahrscheinlich wurde die Spam E-Mail im Auftrag verschickt, da der SMTP-Absender "Return-Path: <n11jsmnp@msn.com>" lautet. Die Hinweise auf das angebliche Opt-In und das angebotene Opt-Out sind für diese Spam E-Mail-Aktion offensichtlich "fake" und führen gewollt in die Irre.

Die Nutzung einer Opt-Out Möglichkeit kann daher zu folgenden Konsequenzen führen:

- Verbirgt sich hinter einem Opt-Out Link eine Web-Seite, können darüber zusätzlich Informationen über den "Besucher" ermittelt werden.
- Wird ein "mailto:" Link angeboten, liegt der Verdacht nahe, dass hierüber eine Adress-Harvesting (Verifikation der Empfänger-Adresse) vorgenommen wird.
- Im günstigsten Fall ist die Opt-Out Adresse ungültig (fake) und dient nur vorgeblich zum Ausschreiben, sowie um die Legitimität der Spam E-Mail zu wahren.

### 7.2.3.3 Joe-Jobs

Wie beschrieben, landen E-Mail-Adressen (z.B. per Harvesting) häufig auf Spam-Listen, ohne dass der potentielle Empfänger dies verhindern kann. Dies ist unangenehm, aber gegen (passiven) den Empfang unerwünschter E-Mails kann

man sich zumindest schützen.

Der umgekehrte Fall ist wesentlich unangenehmer: Die eigene E-Mail-Adresse wird zum Versand von Spam E-Mails missbraucht und somit der (angebliche) Absender diskreditiert. Das sog. Faking von E-Mail-Adressen in der Return-Path Angabe (**Mail From:**) ist trivial; genauso wie die gleiche Adresse im E-Mail Header als "From:" erscheinen zu lassen.

Der erste bekannte Fall dieses Spoofings betraf "Joes.com". Unter [http://ww.spamfaq.net/terminology.shtml#joe\\_job](http://ww.spamfaq.net/terminology.shtml#joe_job) findet sich dazu folgende Erklärung :

*"The act of faking a Spam so that it appears to be from an innocent third party, in order to damage their reputation and possibly to trick their provider into revoking their Internet access. Named after Joes.com, which was victimized in this way by a spammer some years ago."*

Hiergegen gibt es kein probates Mittel. Eine Ausnahme besteht dann, wenn die Spam E-Mail mit der gespoofen Absender-Adresse an Empfänger mit der gleichen Domain-Kennung verschickt wird. Mein SPAMCONTROL-Patch (siehe weiter unten) für Qmail bietet hierzu eine sog. Split-Horizon Überprüfung der SMTP Absender-Adressen.

#### 7.2.3.4 Umfang

Die Frage, welchen Umfang Spam E-Mails mittlerweile ausmachen, muss von unterschiedlichen Perspektiven beleuchtet werden:

- Der Anwender, der ein E-Mail-Konto besitzt, wird die Frage dahingehend beantworten, welchen Anteil Spam E-Mails an seinem gesamten E-Mail-Aufkommen ausmachen.
- Für den Betreiber eines E-Mail-Gateways (= ISP) ist nicht nur die Zahl der zugestellten, sondern auch die die Zahl der unzustellbaren Spam E-Mails relevant. Letzere belasten den E-Mail-Server doppelt: Durch die Annahme der E-Mails und durch das Erstellen einer Bounce-Nachricht an den vermeintlichen Absender ("Return-Path:" im E-Mail Header). Im ungünstigsten Fall antwortet das (unschuldige) Zielsystem — was es nicht tun sollte — oder es wird eine (lokale) Double-Bounce-Nachricht an den Postmaster verschickt: Eine nichtzustellbaren Spam E-Mail generiert also in aller Regel zwei weitere. Im Normalbetrieb kein Problem — bei einer Spam-Attacke aber eine ernste Belastung für den E-Mail-Server, speziell, wenn dieser noch Viren-Checks vorzunehmen hat.

In jedem Fall kostet die Übertragung der Spam E-Mail Bandbreite — häufig zu Lasten des Betreibers des E-Mail-Servers. Glücklicherweise sind Spam E-Mails in der Regel relativ klein; was man bei den heute in HTML/X-HTML verfassten E-

Mails häufig nicht behaupten kann. Wir bedenken, dass nicht nur im Weg des Empfangs von Spam E-Mail Bandbreite vergeudet wird (incoming), sondern auch durch die Erzeugung von Bounce-Nachrichten (outgoing).

Die Grössenverteilung für meine empfangenen Spam E-Mails findet sich in Abbildung 7.2-1; sie folgt im wesentlichen einer Poisson-Verteilung, wobei Spam E-Mails über 100 kByte selten sind und in der Regel darauf hindeuten, dass eine ganze Webseite verschickt wurde.

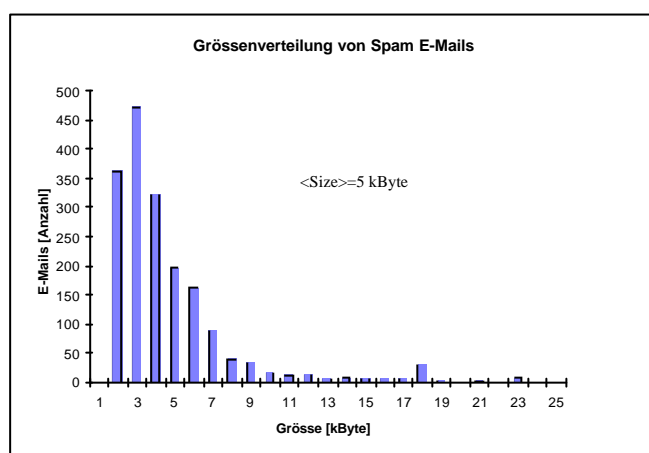


Abbildung 7.2-1: Grössenverteilung der bei mir eingelaufenen Spam E-Mails; der Mittelwert für die Grösse liegt bei etwa 5 kByte.

Zusammengefasst ergibt sich folgendes Bild:

- Die Gesamtzahl der einlaufenden  $N$  Spam E-Mails stellt volumenmässig einen Bruchteil  $X$  der Bandbreite für den (incoming) SMTP-Verkehr dar.
- Von diesen  $N$  Spam E-Mails erreicht ein (kleiner) Teil  $n$  die Mailbox eines Anwenders ( $n/N = 0.1 \dots 0.01$ ).
- $N-n$  E-Mails werden als Bounces an den per Return-Path ("Mail From:") ermittelten Absender bzw. den MTA des ISP zurück geschickt. Hierdurch wird ein Bruchteil Teil  $Y$  der Bandbreite für den (outgoing) SMTP-Verkehr belegt.
- Da die Absenderkennung in der Regel gefälscht ist (*forged*), wird die Annahme der Bounce-Nachricht von einem Teil der MTA mit einem SMTP-Protokollfehler abgelehnt ("No user/mailbox by that name"). Dies resultiert in  $N-n-m$  ( $m < N-n$ ) Double-Bounces an den lokalen Postmaster.
- In ungünstigen Fällen (aufgrund nicht RFC-konformer MTA oder

Autoresponder) ist es möglich, dass auch ein geringer Teil der Bounces zurückgeschickt werden:  $N-n-o$  ( $o \ll N-n$ ).

Es muss strikt unterschieden werden, zwischen dem insgesamt empfangenen Spam E-Mails für eine Domain und den letztlich an eine Mailbox zugestellten (d.h. wirksamen), da — je nach Spam-Aktion — der Spammer beliebige Empfänger in einer Domain adressiert, also quasi mit Schrot schießt.

Insgesamt kann aber das heutige (06/2003) das Spam-Aufkommen und zwar hinsichtlich der Anzahl als auch des Volumes zwischen 20% und 30% des gesamten E-Mail-Verkehrs abgeschätzt werden. Vergleichbare Werte ergeben sich auch aus einer Spam-Analyse von Cord Beermann, wie aus den Abbildungen 7.2-2 und 7.2-3 hervorgeht, und dem ich hiermit für sein Zahlenmaterial danken möchte.

Abhängig davon, wie bekannt (oder einfach erratbar) die E-Mail-Adresse ist, bzw. auf welchen Adresslisten diese gehandelt ist, kann der "persönliche" Spam-Anteil bis 50% oder gar 80% ausmachen. Es ist klar, dass bei solchen Verhältnissen es keine Freude macht, per E-Mail zu kommunizieren.

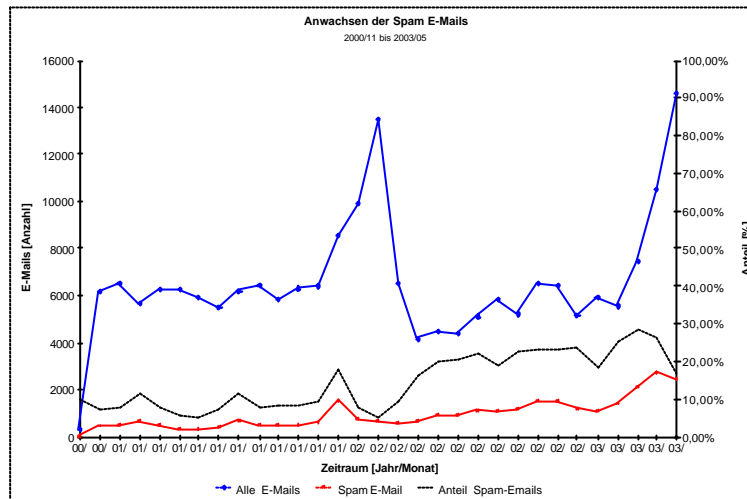


Abbildung 7.2-2: Wachstum der Anzahl der E-Mails und des Spams (Quelle: Cord Beermann).

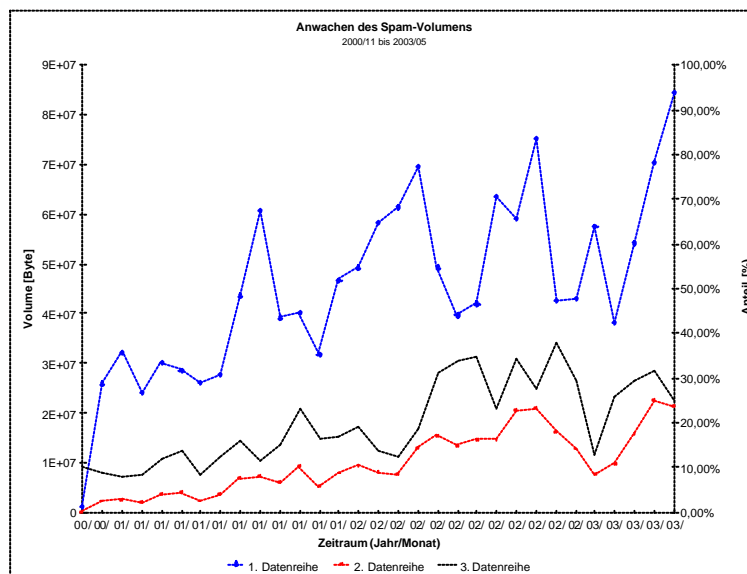


Abbildung 7.2-3: Wachstum des Volumens der E-Mails und des Spams (Quelle: Cord Beermann).

## 7.2.4 Gegenmassnahmen

Je nach Betroffenheitsgrad und Empfänger gibt es verschiedenste Arten der Reaktionen auf Spam E-Mails: Empörung, Ärger, Resignation, störrisches Wegklicken ...

Unbestritten ist, dass mitunter der erhebliche Spam-Anteil an den regulären E-Mails zur Beeinträchtigung des Arbeitsablaufs von Mitarbeitern wird, deren Aufgabe die Bearbeitung und Erledigung von E-Mails ist. Dies betrifft vor allem solche Mitarbeiter bzw. Abteilungen, deren E-Mail-Adresse bekannt, oder relativ leicht erratbar ist. Im besonderen Masse ist dies für öffentliche Institutionen gegeben, deren Auftrag darin besteht, in der (Internet-) Öffentlichkeit Präsent zu sein.

### 7.2.4.1 Rechtliche Massnahmen

In der Presse sind immer wieder Artikel zu lesen, dass sich grosse Firmen (u.a. Microsoft) und die ISPs gegen Spam E-Mails wenden und Musterklagen gegen die Absender dieser E-Mails anstrengen. Der deutsche Gesetzgeber subsumiert Spam E-Mails unter dem allgemeinen Persönlichkeitsrecht und räumt spezielle Regelungen nur Wettbewerbern und Verbänden ein. Wirksam sind Klagen nur dann, wenn der Beklagte im Inland oder in einem über entsprechende bi- bzw. multilaterale Verträge rechtliche justiziablen Ausland befindet. Dies ist in der Regel nicht der Fall.

Hier nutzt auch das diese Tage (20.7.2003) von der Verbraucherministerin angekündigte Gesetzesvorhaben gegen Spam E-Mails nichts. Meines Erachtens gilt für das rechtliche Vorgehen gegen Spam E-Mails und deren Absender der bereits bei der Bewertung von Dienstaufsichtsbewerben zutreffende Spruch: Fristlos, Formlos, Folgenlos.

Internet Service Provider (ISP) müssen sich in der Behandlung der Spam E-Mails auch den Erfordernissen des TKG dar (Telekommunikations Gesetz) sowie des TDSG (Telekommunikation Daten Schutz Gesetz) entsprechen. Sobald Massnahmen des Providers zur Verminderung von Spam E-Mails greifen, macht er sich potentiell der Unterdrückung von Nachrichten schuldig. Ausnahmen hiervon bestehen m.E. jedoch darin, wenn festgestellt wird, dass die E-Mail nicht den gängigen Standards entspricht. Diese Situation ist vergleichbar, wenn ein Brief nicht ordentlich beschriftet oder frankiert ist. Bei Ablehnung erhält der Sender in jedem Fall eine entsprechende SMTP-Mitteilung über die Ursache und kann somit im Zweifelsfall seinen Fehler korrigieren bzw. über den (per E-Mail erreichbaren) Postmaster des Empfängers eine Änderung erwirken.

ISPs und Unterdrückung von  
Nachrichten

### 7.2.4.2 Organisatorische Massnahmen

Jedem von uns flattern mit der normalen Tagespost wöchentlich Dutzende von

Werbesendungen in den Briefkasten ein: Sei es der Pizzabäcker auf Rollschuhen um die Ecke, der Pfarrbrief der Gemeinde, Gewinnspiele, Werbung für Schlankheitskuren, Prospekte der im Umkreis befindlichen Supermärkte ....

Darüber regt sich niemand auf. Gelegentlich liest man beim ein oder anderen Briefkasten "Bitte keine Werbung einwerfen!" — zu einem öffentlichen Thema wird dies nicht, obwohl das Verhältnis von privater Post zu Werbesendungen nicht anders ist als bei Spam zu regulären E-Mails. Warum ist also die Betroffenheit so sehr viel grösser?

Hierfür gibt es m.E. zwei Gründe:

1. Glaube an die Unschuld des Mediums. Bislang wurde das Medium E-Mail als hinreichend sicher und "privat" betrachtet. Versetzen wir uns mehr als 100 Jahre zurück, wo die Post noch per Pferdekutschen ausgetragen wurde, hätte auch niemand akzeptiert, dass hierüber Reklame verteilt würde. Wir haben also ein Problem des Umgangs mit dem neuen Medium.
2. Ein grosser Teil der Inhalt der E-Mails ist sexistischer Natur und teilweise so eindeutig, dass sich viele Menschen in ihrer moralischen Einstellung massiv belästigt und aufgewühlt fühlen — was durchaus beabsichtigtes Ziel der Werbung ist.

Wir haben es also mit einem menschlichen Betroffenheitsproblem zu tun. In grösseren Firmen/Organisationen mag es hilfreich sein, dem strukturiert zu entgegen:

Betroffenheit

- *Schulungen*: Interne Seminare über Spam E-Mails mit den in diesem Abschnitt besprochenen Inhalten. Ziel sollte sein, statt Resignation zu verbreiten, möglichst die "Lacher" auf seiner Seite zu haben.
- *Abuse-Abteilung*: Schulung und Einweisung geeigneter Mitarbeiter zur Bearbeitung und Verfolgung besonders "gemeiner" Spam E-Mails. Dies ist kein einfacher Job, nimmt aber vielleicht den Druck von den Mitarbeitern, die hiervon besonders betroffen sind.
- *Firmeninterne Spam-Info Webseite*: Hier sollte auf das Problem allgemein eingegangen werden, ggf. auf den Spam E-Mail Account und auf die vom Unternehmen getroffenen organisatorisch/rechtlichen Massnahmen hingewiesen werden (u.a. auf die Abuse-Abteilung). Tröstlich für die betroffenen Mitarbeiter könnte auch ein periodisches Veröffentlichen der abgewiesenen Spam E-Mails sein. Ferner sollten noch einmal der "E-Mail-Codex" der Firma verbindlich vorgestellt werden.
- *Anti-Spam E-Mail-Account*: Einrichtung einer interen E-Mail-Adresse, an denen die Mitarbeiter ihre Spam E-Mails weiterleiten können und damit sie diese alsdann löschen (z.B. über ein Regelwerk). Die eingesammelten Spam E-

Mails können analysiert und ggf. über die Abuse- oder Rechtsabteilung weiter verarbeitet werden.

- *Öffentliche Anti-Spam Webseite:* Bei aller Vorsicht und Sorgfalt kann es doch mitunter passieren, dass eine eigentlich unverdächtige und mitunter sogar wichtige E-Mail aufgrund der getroffenen Anti-Spam-Verfahren nicht angenommen wird und der Absender nach Möglichkeiten sucht, ggf. von einer Blocking List gestrichen zu werden. Hierzu ist es hilfreich (eine nicht verlinkte) Anti-Spam Webseite als URL zu referenzieren, die bei der Ablehnung auf IP- oder SMTP-Niveau dem vermeintlichen Spam-Sender mitgeteilt wird (vgl. **rbl dns** und **tcpserver** weiter unten). Auf dieser Webseite sollten Kontaktinformationen veröffentlicht sowie die Eingabe von "Complaints" mittels HTML-Forms ermöglicht werden.

#### 7.2.4.3 Vermeidung

Einige weitere wichtige Punkte ergibt sich hinsichtlich der Vermeidung von Spam E-Mails:

- Hilfreich ist sicherlich, den Mitarbeitern nicht nur ein E-Mail Konto, sondern mehrere getrennte zur Verfügung zu stellen, z.B. auch eines zur Abwicklung der privaten Korrespondenz. Sofern sich die Mitarbeiter in öffentlichen Diskussionsformen engagieren, sollte darauf geachtet werden, nicht die firmen-eigene E-Mail-Adresse dort zu nutzen bzw. bekanntzugeben. Dies erlaubt in jedem Fall ein besseres Filtern von evtl. Spam E-Mails.
- Den Mitarbeitern sollte klar gemacht werden, dass auch sie nicht aktiv zur Weitergabe von E-Mail-Adressen beitragen sollten. Hierzu zählt vor allem der Verzicht auf die Angabe mehrerer Empfänger von E-Mails in der "To:" oder "CC:" Zeile. Sollen dennoch mehrere Adressaten erreicht werden, lassen sich diese per "BCC:" (*Blind Carbon Copy*) einfügen.
- Bei E-Mail-Adressen, die per Web erreicht werden können, bietet es sich an, mit HTML-Forms zu arbeiten. Sollte aber dennoch eine E-Mail-Adresse per "mailto:" bekannt gemacht werden, kann diese z.B. periodisch wechseln, beispielsweise per "adresse-YYYYMM", indem ein Zeitstempel eingefügt wird. Qmail erlaubt durch die Vergabe von VERP (*Variable Envelope Response Path*) ein einfach steuerbares Filtern dieser Adressen; ohne dass der Absender sofort eine Bounce-Nachricht erhält.

#### 7.2.4.4 Technische Möglichkeiten

In der Praxis haben sich mehrere Ansätze etabliert, der Flut der Spam E-Mails entgegenzutreten:

- *Blockieren*, d.h. die E-Mails bereits zum Zeitpunkt der Verbindungsaufnahme entweder auf IP-Niveau oder auf SMTP-

(=Anwendungs)-Ebene zu blockieren, sodass keine eigentlichen Nutzdaten übertragen wurden.

- *Filtern*, d.h. Empfang aller E-Mails und anschließende Analyse der E-Mails, nach dem alten Verfahren: "Die Guten ins Töpfchen, die Schlechten ins Kröpfchen". Hierbei hat man die Wahl, auf die Schlechten — also als Spam identifizierten E-Mails — anschließend zu reagieren. Soll die Spam E-Mail allerdings erhalten bleiben, setzt dieses Verfahren einen lokalen MUA mit mehreren adressierbaren Mailboxen (oder Mailverzeichnisse) voraus, was beim entfernten Zugriff über IMAP4 realisiert werden kann.
- *Tagging* zeichnet die vorliegende Spam-Merkmale (bzw. den sog. Spam-Level bzw. Spam-Score) in einem eigenen E-Mail-Header auf und erlaubt es somit dem Anwender selbst mittels eigener Filterregeln auf die identifizieren Spam E-Mails zu reagieren.
- *Digesting* funktioniert in Abgrenzung zu den anderen Verfahren lediglich auf User-Basis. Der E-Mail-User führt hierbei normalerweise eine sog. *Whitelist* von Absenderadressen, von denen ohne Einschränkung E-Mails angenommen werden. Jeder andere Sender muss sich zunächst durch eine zusätzliche Reply-E-Mail in Form eines *Message Digest* qualifizieren, um auf diese Whitelist zu gelangen.

In der Praxis wird häufig eine Mixtur dieser Verfahren angewandt, um ihre Stärken zu kombinieren. Die Problematik, die sich beim Einsatz jeder dieser Methode prinzipiell ergibt, lautet:

- *False Positives*: Reguläre E-Mails werden als Spam eingeschätzt und ggf. verworfen bzw. abgelehnt.
- *False Negatives*: Spam E-Mails werden als normale E-Mails betrachtet und gelangen trotz aller Massnahmen zum Adressaten.

Letzteres ist eine Frage der Effizienz; erstere eine der Schwellwerte und des Verfahrens.

#### 7.2.4.5 Was sagen die RFC?

Obwohl Spam beileibe kein neues Phänomen ist, gibt es in der Sammlung der RFC kaum geeignete Hinweise in Form von "Best Current Practice (BCP)". Der RFC 2505 "Anti-Spam Recommendations" gibt allerdings einige Empfehlungen zum Umgang mit Spam E-Mails und weist im besonderen folgende Anforderungen an MTAs aus:

- 1) *MUST be able to restrict unauthorized use as Mail Relay.*
- 2) *MUST be able to provide "Received:" lines with enough information to make it possible to trace the mail path, despite*

*spammers use forged host names in HELO statements etc.*

- 3) *MUST be able to provide local log information that makes it possible to trace the event afterwards.*
- 4) *SHOULD be able to log all occurrences of anti-relay/anti-spam actions.*
- 5) *SHOULD be able to refuse mail from a host or a group of hosts.*
  - 6a) *MUST NOT refuse "MAIL From: <>".*
  - 6b) *MUST NOT refuse "MAIL From: <user@my.local.dom.ain>".*
  - 7a) *SHOULD be able to refuse mail from a specific "MAIL From:" user, <foo@domain.example>.*
  - 7b) *SHOULD be able to refuse mail from an entire "MAIL From:" domain <.\*@domain.example>.*
- 8) *SHOULD be able to limit ("Rate Control") mail flow.*
- 9) *SHOULD be able to verify "MAIL From:" domain (using DNS or other means).*
- 10) *SHOULD be able to verify <local-part> in outgoing mail.*
- 11) *SHOULD be able to control SMTP VRFY and EXPN.*
- 12) *SHOULD be able to control SMTP ETRN.*
- 13) *MUST be able to configure to provide different Return Codes for different rules (e.g. 451 Temp Fail vs 550 Fatal Error).*

In Abschnitt 1.5 des RFC wird die Frage gestellt "Where to block spam, in SMTP, in RFC822 or in the UA", endet aber lediglich mit dem Verweis, dass Spam E-Mails während des SMTP-Dialogs abzuweisen sind.

Leider ist die Situation in der Praxis noch vertrackter:

- Aus den SMTP-RFC lassen sich keine (direkten) Berechtigungen ableiten, E-Mails aufgrund irgendwelcher Kriterien zu blockieren, ausser für unberechtigtes Relaying. Dies betrifft insbesondere den Return-Path ("Mail From:") und die HELO/EHLO-Begrüssung des SMTP-Client. Ferner gilt diese Aussage für den SMTP-Envelope, den E-Mail-Header, als auch natürlich für den E-Mail-Body. Entsprechende Formulierungen in den RFC 2821/2822 lauten in der Regel "Should".
- Im Gegenzug, gibt es keine "Must" oder "Shall" Kriterien, auf die die E-Mail des Absenders überprüft werden sollte. Der SMTP-Server muss so tolerant wie möglich sein. Qmail z.B. lehnt aus sich heraus nur E-Mails mit einem einfachen LF (statt CRLF) ab.

- Beim Einsatz "scharfer" Blockade-Massnahmen für Spam, kann es leicht passieren, dass — aufgrund der "False Negatives" — der Provider auf der Liste der RFC-Ignorant ([www.rfc-ignorant.org](http://www.rfc-ignorant.org)) gesetzt wird; kein sonderliches Vergnügen und auf jedenfalls der Reputation schädlich.
- Die Blockade-Massnahmen bewegen sich somit in einer Grauzone; dies ist wohl mit ein Grund, warum einige grosse Freemail-Anbieter sich auf das Filtern von Spam E-Mails beschränken.

Insgesamt ist die Internet-Community ziemlich unschlussig, wie mit Spam E-Mails umzugehen ist. Die Bandbreite schwankt zwischen grosser Toleranz beim Empfang selbst (SMTP-)syntaktische unzulänglicher E-Mails bis zu einer Ablehnung der Empfang von E-Mails von Dial-Up SMTP-Clients. RFC 1123 "Requirements for Internet Hosts -- Application and Support" sagt z.B. hinsichtlich des HELO Parameters (5.2.6):

*"The sender-SMTP MUST ensure that the <domain> parameter in a HELO command is a valid principal host domain name for the client host. As a result, the receiver-SMTP will not have to perform MX resolution on this name in order to validate the HELO parameter. The HELO receiver MAY verify that the HELO parameter really corresponds to the IP address of the sender. However, the receiver MUST NOT refuse to accept a message, even if the sender's HELO command fails verification.*

*DISCUSSION:*

*Verifying the HELO parameter requires a domain name lookup and may therefore take considerable time. An alternative tool for tracking bogus mail sources is suggested below (see "DATA Command"). Note also that the HELO argument is still required to have valid <domain> syntax, since it will appear in a Received: line; otherwise, a 501 error is to be sent."*

Dieses "MUST" Kriterium für SMTP-Clients ist unter den heutigen Bedingungen von vorwiegend temporären Internet-Dial-Up Verbindungen schlicht hinfällig.

### 7.2.5 Spam E-Mails im Verarbeitungszyklus

Zur Entscheidung ob eine E-Mail Spam ist oder nicht, bedarf natürlich hinreichender Informationen. Je nach Kenntnisstand kann die Spam E-Mail zu unterschiedlichen Verarbeitungszeitpunkten sehr differenziert behandelt werden, was nachfolgende Tabelle in einem Gesamtüberblick liefert:

| Zeitpunkt | Schicht/(Schritt) | Informationen | Massnahmen |
|-----------|-------------------|---------------|------------|
|-----------|-------------------|---------------|------------|

| Zeitpunkt                             | Schicht/(Schritt)      | Informationen   | Massnahmen  |
|---------------------------------------|------------------------|---|---|
| Aufbau der TCP/IP-Verbindung          | IP (Layer 3)           | Sender-IP,<br>FQDN des Senders,<br>Blocking List  | Blockieren<br>[per MTA]   |
| SMTP-Dialog                           | SMTP (Layer 7)         | Return-Path (MAIL From:),<br>Forwarding-Path (RCPT<br>To:)<br>MX der Return-Path<br>Domain,<br>HELO/EHLO Angabe | Blockieren (SMTP<br>Fehlercode 5xx),<br>Deferral (Verzögern,<br>SMTP Fehlercode<br>4xx)<br>[per MTA]        |
| Einfügen in die Queue                 | (Verarbeitung)         | E-Mail-Header,<br>E-Mail-Body (Inhalt/Text),<br>MIME-Struktur,<br>Footprint (remote)                            | Bounce,<br>Verwerfen,<br>Taggen<br>[per MTA]  |
| Ausliefern in Empfänger-Mailbox       | (lokale Zustellung)    | E-Mail-Header,<br>E-Mail-Body (Inhalt/Text),<br>MIME-Struktur,<br>Footprint (remote);<br>evtl. Spam-Tags        | Bounce,<br>Verwerfen,<br>Ablage "Junkmail"<br>Ordner,<br>Taggen<br>[per User]                               |
| Abholung durch Remote Mail User Agent | (entfernte Zustellung) | E-Mail-Header,<br>E-Mail-Body (Inhalt/Text),<br>MIME-Struktur,<br>Footprint (remote);<br>evtl. Spam-Tags        | Abholen,<br>Nicht-Abholen,<br>Löschen nach <i>N</i><br>Tagen,<br>lokaler "Junkmail"<br>Ordner<br>[per User] |

Tabelle 7.2-1: Informationen über E-Mails und Möglichkeiten zur Behandlung von Spam E-Mails im Verarbeitungszyklus.

Bei der Ablehnung von E-Mails auf IP-Schicht oder im Laufe des SMTP-Dialogs handelt es sich mithin um eine **Blockade**. Ist die E-Mail einmal in Empfang genommen, kann sie nach unterschiedlichen Kriterien einem oder mehreren **Filtern** unterworfen werden. Dies resultiert in einer — je nach Resultat — unterschiedlichen Verarbeitung der E-Mails (*aktives Filter*). Im Gegensatz hierzu wird beim **Tagging** lediglich das Filterergebnis beispielsweise in der zugefügten E-Mail Header-Zeile vermerkt (*passives Filter*) und die Aufgabe des aktiven Filterns den anschliessend ablaufenden Programmen überlassen.

Beachtenswert ist insbesondere auch, dass das Blockieren auf Grundlage MTA-

---

spezifischer Regeln erfolgt, währenddessen das Filtern sowohl "global", d.h. MTA- wie auch benutzerspezifisch vorgenommen werden kann.

Exemplarisch ist der Verarbeitungszyklus für Qmail in Abbildung 7.2-4 dargestellt. Zusätzlich ist es möglich, die auf Schicht 3 (IP) und 7 (SMTP) ebenfalls in einem zusätzlich eingefügten E-Mail-Header Feld zu hinterlegen, und so für die spätere Auswertung verfügbar zu machen.

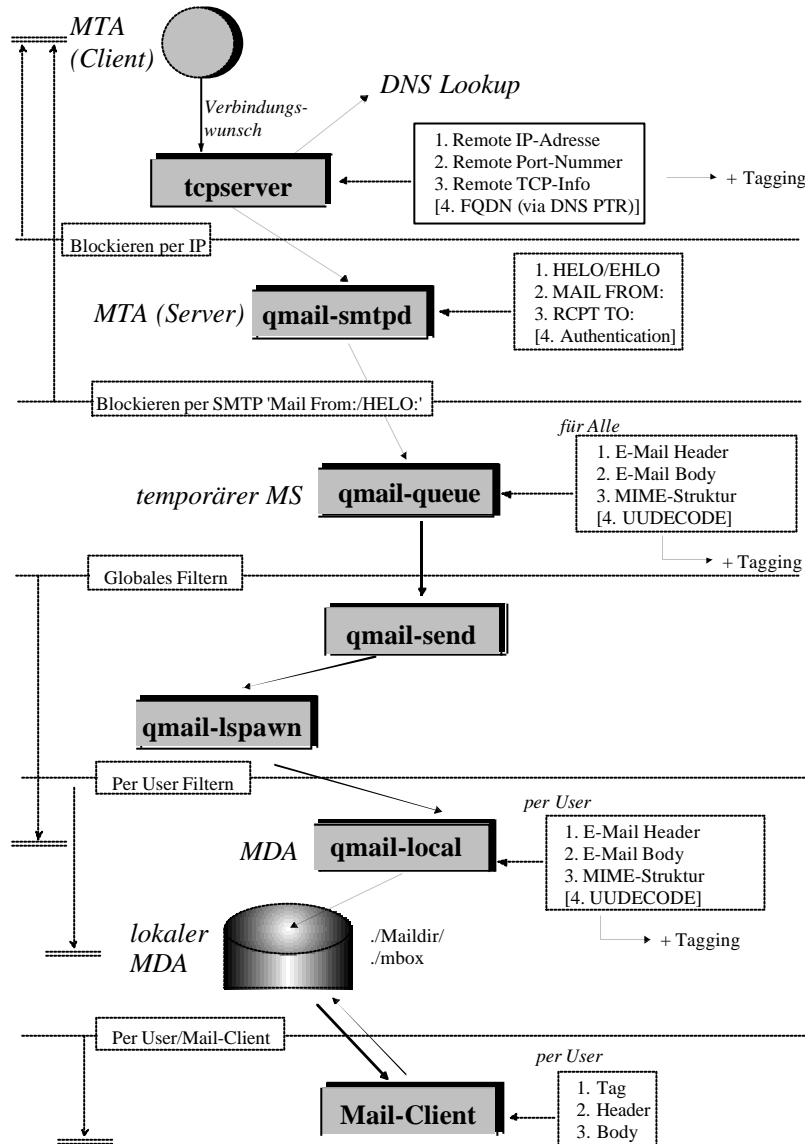


Abbildung 7.2-4: Mögliche Schritte der Verarbeitung von Spam E-Mails bei Qmail.

### 7.2.5.1 Blockieren

Das Blockieren von E-Mails kann auf IP-Schicht und/oder auf SMTP-Schicht

erfolgen; wem das OSI-Referenzmodell vertraut ist, wird die IP-Schicht mit dem Netzwerk (Schicht 3) und SMTP mit der Anwendungsschicht (Schicht 7) identifizieren. Folgende Vorteile ergeben sich beim Blockieren von Spam E-Mails:

- Beim Blockieren einer E-Mail erhält der Absender in der Regel eine qualifizierte Mitteilung hinsichtlich der Ablehnung.
- Das Verfahren ist unabhängig davon, ob der Empfänger auch tatsächlich existiert, da es ausschliesslich Absender-spezifisch erfolgt.
- Da nur sehr wenig Protokollinformation zur Feststellung benötigt wird, erfolgt die Ablehnung zum frühest möglichen Zeitpunkt; es wird nicht die eigentliche Mitteilung übertragen und somit Bandbreite gespart.

Das Verfahren hat aber auch einige Nachteile:

- Das Blockieren arbeitet MTA- und nicht Benutzer- (=Empfänger-)spezifisch.
- Erfolgt die Ablehnung bereits auf IP-Niveau, können auch legitime Absender davon betroffen sein. So haben es z.B. einige Firmen zu eigen gemacht, keine E-Mails von Absendern mit Dial-Up IP-Adressen in Empfang zu nehmen. Um zu diesen E-Mails zu übertragen, muss man — per SMTPROUTE — notgedrungen über das E-Mail-Relay des Providers gehen.
- Es kann nicht garantiert werden, dass der initiale Sender die Information über die Ablehnung erhält. Dies kann dann der Fall sein, wenn er über ein SMTP-Relay seine E-Mail verschickt.

Der effizienteste Weg zur Abwehr von Spam E-Mails ist deren Ablehnung auf IP-Niveau bei Verbindungsaufnahme:

tcpserver

Sender (IP; TCP-Port) ==> Empfänger (IP; TCP-Port=25)

Zudem wird häufig über einen DNS-Lookup (`tcpserver -h`) während des Verbindungsversuchs der FQDN des Sender ermittelt (A Record). Es ist möglich, die Verbindungsaufnahme mittels `tcpserver` unter den folgenden Szenarien zu blockieren:

- Die IP-Adresse, bzw. der IP-Adressbereich wird blockiert:

```
61.255.105.123:deny # deny, if from that IP
210.10.10.:deny    # deny, if form that net
```

- Mit dem `tcpserver` Flag "Paranoid" lässt sich ferner die Konsistenz von A- und PTR-RR überprüfen (`tcpserver -p`), indem zusätzlich über den FQDN und die Abfrage des PTR-Records erneut die IP-Adresse ermittelt wird. Bei einem Mismatch wird dann die Environment-Variable `$TCPREMOTEHOST` gelöscht (bei nicht feststellbaren FQDN stellt Qmail dann den Namen als "unknown" dar). Das `tcpserver`-Regelwerk hierzu lautet:

```

=:allow # accept, if $TCPREMOTEHOST is set
:deny # deny all other connections

```

- Die IP-Adresse ist eine temporäre Dial-Up Adresse; der Provider mappt diese häufig unter dem Namen "dial" oder mittels ihres Reverse-Namens. T-Online Dial-Up Sender und solche mit einem Reverse-Namen können über das **tcpserver**-Regelwerk wie folgt abgewiesen werden:

```

=.in-addr.arpa:deny # disable, if reverse IP name
=.t-dialin-net:deny # disable, if from t-online

```

Mittels des Gleichheitszeichens "=" wird beim Aufbau bzw. Auswertung der **tcpserver** cdb kenntlich gemacht, dass nicht die IP-Adresse (via der Environment-Variablen **\$TCPREMOTEIP**), sondern der FQDN (entsprechend **\$TCPREMOTEHOST**) genommen werden soll, wobei im Fall der IP-Adresse die Auswertung von Rechts-nach-Links und bei den Namen von Links-nach-Rechts erfolgt.

- Die mögliche Ausgabe eines sog. *Banner* durch **tcpserver** bei der Verbindungsaufnahme (**tcpserver -B "No Spam"**) ist in der Regel nicht hilfreich, um dem SMTP-Client einen Hinweis auf seine Ablehnung zu geben; dies leistet aber — wie weiter unten dargestellt — **rblsmtpd** zusammen mit **tcpserver**.

Soweit mir bekannt, stammt die Idee, die Entgegennahmen von E-Mails zu verweigern, die von "offenen" MTAs kommen, oder aktiv an der Weitergabe von Spam E-Mails beteiligt sind von Paul Vixie. Dieser rief Mitte der 90'er Jahre das Projekt *Mail Abuse Protection System* MAPS ins Leben, in dem solche MTAs aufgelistet sind und in Echtzeit— während der Mailverarbeitung per DNS-Lookup — ausgewertet werden können, was als *Real-time Blocking List* (RBL) bezeichnet wird.

Real-time Blocking List  
(RBL)

MAPS ist zwischenzeitlich nicht mehr aktiv und Organisation wie ORBS, SpamCop, ORDB sowie viele andere realisieren heute vergleichbare Dienste. Der Name MAPS wird zwischenzeitlich vom Nachfolger *Mail Abuse Prevention System LLC* MAPS<sup>SM</sup> (<http://mail-abuse.org/>) verwendet. Alle RBL-Anbieter gemeinsam ist, dass sie sich in einer Grauzone des Internet befinden. Niemand kann garantieren, dass die aufgenommen Server auch wirklich offene Relays sind, der Test vielleicht nicht sinnvoll ausfällt oder gar der Anbieter auf einen "Joe-Jobs" hereingefallen ist. Firmen — speziell Internet-Anbieter und FreeMailer — führen daher mitunter einen rechtlichen Kampf gegen die Blocking List-Betreiber, mit der Folge, dass die Dienste schnell verschwinden und genauso schnell unter neuer Bezeichnung und Lokation ihren Dienst wieder aufnehmen.

Der "Scope" der Blocklisten-Betreiber hat sich mittlerweile enorm erweitert und betrifft nicht mehr ausschliesslich den SMTP-Datenverkehr. Der RBL-Dienst kann

effektiv von beliebigen TCP/IP-Anwendungen genutzt sind. So finden sich in einigen Blocklisten ganze Netzwerke, die somit vom Internet-Datenverkehr "gebannt" werden können.

Eine Spielart der Realtime Blocking List stellt die DUL (*Dial Up User List*) Datenbank dar. Während in der RBL "bekannte" MTAs mit festen Adressen auftauchen, landen in der DUL die von Providern dynamisch vergebenen IP-Adressen. Hinter den eingetragenen IP-Adressen finden sich nicht notwendigerweise Spam-Sender; der Nutzer des DUL-Dienstes sieht aber diese IP-Adressen als "unerwünscht" an, um eine SMTP-Verbindung mit seinem Server aufzunehmen.

**Dial Up User List (DUL)**

Wie bei allen RBL-Diensten gilt es auch bei der Nutzung der DUL, den Anbieter sehr genau zu eruieren und sich mit dessen "Policy" vertraut zu mache; speziell wenn Problemfälle auftreten.

Das Verfahren, nachdem die RBL aufgebaut sind, ist sehr effizient: Der Betreiber der Blocking List stellt einen *Authoritive Domain-Server* für eine High-Level Domain zur Verfügung, auf dem ein spezieller Name-Server läuft (z.B. **rblDNS** aus dem DJBDNS-Package von Dan Bernstein).

**Blocking List-Server**

- Jedes erkannte offene Relay wird über einen eigenen DNS TXT-Record (RFC 1035) eingetragen, indem die umgekehrte [a.b.c.d => d.c.b.a] IP-Adresse mit dem FQDN des Blocking List-Servers konkatiniert und unter diesem veröffentlicht wird. Inhalt des TXT-Records im positiven Fall üblicherweise ein erklärender Text sowie ein Verweis (URL) auf die Webseite des Blocking List-Betreibers; liegt kein entsprechender TXT-Record vor, bleibt die Antwort "leer".
- Wird nun ein DNS TXT-Lookup (z.B. mit dem Programm **dnstxt** aus DJBDNS) auf diesen Namen vorgenommen, erhält man eine entsprechende Antwort; z.B. für den ORBD-Diensts mit dem FQDN [relays.ordb.org](http://relays.ordb.org) und der Testadresse 127.0.0.2:

```
# dnstxt 2.0.0.127.relays.ordb.org
Listed by ORDB - for testing purposes only
```

Diese Vorgehen kann als *Pre-Initial-Spam-Verfahren* bezeichnet werden, wobei in jedem Fall ein spezielles Dienstprogramm wie **rblsmtpd** eingesetzt werden muss, auf das im Anschluss eingegangen wird.

Die Betreiber der Blocking List-Server (vgl. <http://www.geocities.com/spamresources/filter-dnsbl.htm>) haben unterschiedliche Strategien:

- **Scanning**: Die im Internet per MX-Record eingetragenen Mail-Server werden daraufhin überprüft, ob sie sich ggf. wie ein offenes Relay verhalten.

- **Testing:** Über eine Web-Interface des Betreibers ist es in der Regel möglich, selbst z.B. den eigenen MTA zu testen.
- **Aufnahme:** Über Complaints von Anwendern können Mailserver auf diese Liste gesetzt werden.
- **Löschen:** Nach dem Absichern eines MTA muss es natürlich möglich sein, den Eintrag innerhalb einer kurzen Frist wieder zu entfernen.

Mittels des Programms **rblsmtp** aus dem UCSPI-Paket von Dan Bernstein kann ein Lookup mehrerer RBL-Listen vorgenommen werden. Ein typischer Aufruf unter **tcpserver** und beim Einsatz von Qmail lautet (vgl. Abbildung 7.2-5):

rblsmtpd

```
exec softlimit -m 2000000 \  
    tcpserver -vRh -l $HOSTNAME \  
    -x /var/qmail/etc/locals.cdb \  
    -u $QMAILDUID -g $QMAILDGID 0 smtp \  
    rblsmtpd -C -b -r relays.ordb.org \  
    /var/qmail/bin/qmail-smtpd 2>&1
```

Mit dem Aufruf **rblsmtpd -C -B -r relays.ordb.org** wird folgendes erreicht:

- **-C:** Fail-Open Mode; falls der Dienst im DNS temporär nicht erreicht ist, wird die Verbindung nicht abgewiesen (ansonsten **-c**),
- **-b:** Ausgabe des SMTP-Fehler Codes '553' (permanente Ablehnung) statt wie per Default (**-B**) '451', d.h. lediglich deferral.
- **-r relays.ordb.org:** Angabe des FQDN der RBL-Quelle, wobei sich mehrere Quellen auflisten lassen (falls **-a FQDN** eintragen wird, handelt es sich nicht um eine Blocking List, sondern um eine Whitelist).

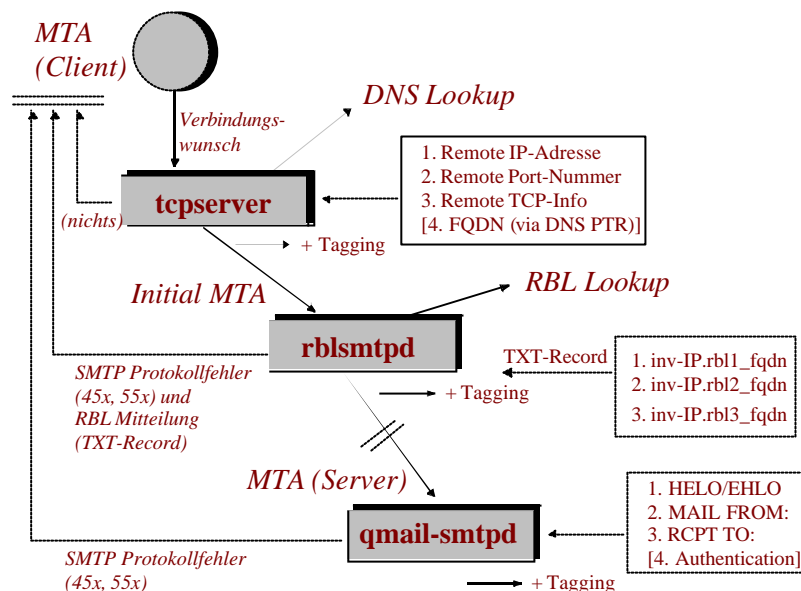


Abbildung 7.2-5: Zusammenspiel von **tcpserver**, **rblsmtpd** und **qmail-smtpd**.

**rblsmtpd** verhält sich wie ein "normaler" SMTP-Server, d.h. er führt den ersten Schritt des SMTP-Dialogs durch, indem er dem Client zunächst mit **220 rblsmtpd.local** antwortet und parallel hierzu einen DNS-Lookup in der angegebenen Blocking List vornimmt. Per Default, d.h. ohne Angabe einer expliziten RBL-Base, nimmt **rblsmtpd** einen Lookup bei <http://maps.vix.com/rbl/> vor; einer nicht mehr existenten RBL.

Abhängig vom Resultat des Lookups verhält sich **rblsmtpd** wie folgt:

- Ist die Antwort negativ — d.h. der Sender ist nicht in der RBL vorhanden — wird die Environment-Variable **\$RBLSMTPD** nicht gesetzt und **rblsmtpd** startet nun den eigentlichen SMTP-Sever, z.B. **qmail-smtpd**.
- Ist die Antwort positiv wird die Environment-Variable **\$RBLSMTPD** gesetzt, bzw. mit der Antwort der RBL befüllt. **rblsmtpd** führt anschliessend den SMTP-Dialog bis zum "RCPT To: <Forwarding-Path>" fort und gibt erst dann einen SMTP-Fehlercode zurück. Hierbei wird die Information dem SMTP-Client zurückgegeben, die als Inhalt des DNS-TXT Records eingetragen ist bzw. nun in **\$RBLSMTPD** steht. Beispiel:

```
220 rblsmtpd.local
helo me
250 rblsmtpd.local
```

```
mail from: <spamtest@example.com>
250 rblsmtpd.local
rcpt to: <erwin>
451 Blacklisted IP address 192.168.192.3; see
http://www.fehcom.de/qmail/spaminfo.html
```

- Ein Spezialfall liegt vor, wenn der Inhalt der Environment-Variablen `$RBLSMTPD` mit einem Bindestrich "-" beginnt. In diesem Fall gibt **rblsmtpd** immer eine '553' SMTP-Fehlermeldung an den Client zurück, sodass der Absender eine sofortige Bounce-Nachricht erhält.

Eine elegante Möglichkeit, sich der vermeintlichen Autorität von RBL-Betreibern zu entsagen, ist der Aufbau eines eigenen (unternehmensweiten) RBL-Servers. Notwendig hierzu sind folgende Instrumente:

Lokale RBL

1. Eine eigener lokaler RBL-Dienste, z.B. der bereits erwähnte **rbl dns** von Dan Bernstein. Dieser muss als Nameserver im lokalen DNS den Clients (=MTAs, die auf diesen Dienst zugreifen) im allgemeinen DNS der Firma bekannt gemacht werden.
2. MTAs, die z.B. via **tcpserver** und **rblsmtpd** auf die Dienste dieses lokalen **rbl dns** lauschen.
3. Eine möglichst automatische Registrierung von Spam E-Mails mit der Möglichkeit, die IP-Adresse des Absenders in die **rbl dns**-Datenbank unverzüglich einzustellen.
4. Gegebenenfalls kann es sich als nützlich erweisen, für das Unternehmen eine eigene "Abuse" Web-Seite zu veröffentlichen, die in der **rblsmtpd**-Mitteilung als URL referenziert wird.

In der Regel läuft der **rbl dns** auf dem MTA, auf dem auch die Spam-Analyse durchgeführt wird, wodurch sich dieses Konstrukt auch für alleinstehende Server eignet. Die Arbeitsweise eines lokalen RBL-Dienstes ist verblüffend einfach:

Ist eine eingelaufene Spam E-Mail als solche erkannt, erfolgt die sofortige Aufnahme der IP-Adresse des sendenden MTAs in die RBL, sodass alle weiteren E-Mails von diesem Absender (bis zum Zeitpunkt eines möglichen Entfernens — *Agings*) blockiert werden. Von den in einer typischen Kampagne einlaufenden mehreren tausend Spam E-Mails erreicht somit im Idealfall nur eine einzige ihr Ziel; alle anderen werden abgewiesen. Kombinieren kann man diese Methode mit speziell präparierten E-Mail-Adressen auf den öffentlichen Web-Seiten der Firma, die dann als funktionales Spam-Ziel fungiert.

Dieses *Post-Initial-Spam-Verfahren* ist recht effektiv und zuverlässig. Da **rblsmtpd** zudem eine sinnvolle Mitteilung an den sendenden MTA abgegeben kann, ist diese Methode auch ziemlich sicher hinsichtlich "False Negatives". Wird

zudem ein Aging der IP-Adressen eingesetzt und als SMTP-Fehlercode — wie per Default beim **rbldsmtpd** vorgesehen — ein SMTP '451' Fehlercode ausgegeben, haben unschuldig betroffene Absender kein Nachsehen, da mittels der üblichen Queuing-Massnahmen die E-Mail zu einem späteren Zeitpunkt erfolgreich zugestellt werden kann, oder aber der betroffene Absender im schlechtesten Fall eine "böse" E-Mail an den Postmaster (z.B. per Abuse Web-Seite mitgeteilt) verschickt (dessen E-Mail-Adresse sich natürlich dann häufig ändern sollte).

Das sicherlich gängigste Mittel zur Spam-Bekämpfung ist die Blockieren auf SMTP-Niveau durch eine der folgenden Massnahmen:

SMTP

- Der SMTP-Absender ist auf einer Blocking Liste — bei Qmail **badmailfrom**.
- Der SMTP-Absender (genauer: die Return-Path Angabe im "Mail From: <Return-Path>") kann per DNS nicht aufgelöst werden; entweder weil der Domain-Teil der Adresse falsch geschrieben wurde, nicht-zulässige Zeichen enthält (z.B. Leerzeichen), die Domain nicht existiert oder für die Domain kein MX-Eintrag vorliegt, sodass ein evtl. Bounce nicht zustellbar sind.
- Im HELO/EHLO-Statement wird keine oder keine gültige FQDN-Adresse angegeben, bzw. diese kann via DNS-Lookup nicht ermittelt werden.
- Als Tarptitting wird eine zusätzliche Massnahme bezeichnet, die dann greift, wenn der Absender E-Mails an mehrere Absender "RCPT To:" absetzen will. Hierbei wird ein Zähler geführt der bei Überschreitung eine Verzögerung (*sleep*) in der Entgegennahme des nächsten "RCPT To:" einführt. Hierdurch wird der Absender gezwungen zu warten und es erscheint ihm, als wäre er in ein "Teerfass" gefallen.

Diese Verfahren funktionieren sehr gut, sofern die Spam-Erzeuger/Sender das SMTP-Protokoll nicht richtig verstanden haben. Häufig wird z.B. in Ermangelung einer geeigneten HELO/EHLO-Kennung die (einfach zu ermittelnde) IP-Adresse des Empfangssystems als SMTP-Begrüssung benutzt: Bingo.

Allerdings wird keines dieser praktikablen Verfahren von gängigen RFCs gedeckt; sie sind im Grunde nicht zulässig. Dies gilt umso mehr, wenn im Zuge eines auf diese Weise dem Spam-Sender zugetragenen SMTP '5xx' Fehlercodes zusätzlich die Verbindung gekappt wird.

Das von mir bereitgestellte SPAMCONTROL-Patch (<http://www.fehcom.de/qmail/spamcontrol.html>) für Qmail vereinigt die meisten der gängigen Anti-Spam-Verfahren auf SMTP-Niveau. Zudem ermöglicht es ein qualifiziertes Loggen, falls eine Filtermassnahme greift und kann so sehr schnell zur Kontrolle der Spam-Aktivität bzw. der Effizienz eingesetzt werden:

SPAMCONTROL

- Filtern auf die "MAIL From:" (Return-Path) Adresse mit Wildcards.
- Filtern auf die "RCPT To: <Forwarding-Path>" Adresse mit Wildcards.
- Filtern auf die HELO/EHLO Angabe mit Wildcards.
- Überprüfung des Domainteils in der Absenderangabe "MAIL From: <Return-Path>" im DNS.
- Überprüfung der Hostangabe in der HELO/EHLO Mitteilung im DNS.
- Tarpitting mit variablen Grenzwerten.

Als weiteres nützliches Feature arbeiten die meisten Filter im *Split-Horizon* Verfahren. D.h. es wird eine Fallunterscheidung gemacht, ob ein Absender von einem vertrauenswürdigen Sende-Host die E-Mail verschickt oder nicht. Bei Qmail macht sich dies durch das Setzen der Environment-Variablen \$RELAYCLIENT fest; ist diese gesetzt kann der Absender E-Mails beliebig verschicken, ansonsten nur an die per `rcpthosts/morercptshots` deklarierten Domains. Hierdurch werden (intern wie extern) E-Mails mit gespoofen Absenderadressen erkannt und ihre Weiterleitung unterbunden.

#### 7.2.5.2 Strukturelle Analyse

Zur strukturellen Analyse von E-Mails müssen diese daher natürlich zunächst angenommen worden sein. Die Analyse kann zu folgenden Zeitpunkten stattfinden:

1. Unmittelbar nach dem Entgegennahme durch den SMTP-Server (**qmail-smtpd**) on-the-Fly beim Einstellen in die Queue.
2. Während der lokalen Zustellung durch den Mail Delivery Agent MDA (**qmail-local**).
3. Nach dem Herunterladen der E-Mail über das Protokoll POP3/IMAP4 als Plug-In für den lokalen Mail User Agent MUA (Outlook, The Bat, Eudora, Netscape ...).

Die ersten beiden Verfahren arbeiten also Empfangs-System basiert und sind stets dann notwendig, wenn der Empfänger seine E-Mail nicht ausschliesslich über einen lokalen MUA, sondern speziell über ein Web-Interface Zugriff erhält. Alle potentiellen E-Mail Empfänger werden identisch behandelt, d.h. die strukturelle Analyse erfolgt mittels einer einheitlichen Engine (MTA-basierend), wobei im zweiten Fall zusätzlich benutzerspezifische Kriterien einfließen können.

Das letzte Verfahren ist natürlich am Ressourcen-schonendsten, da die Analyse nun auf den Clients stattfinden. Viele E-Mail Programme bieten Plug-Ins für Anti-Spam Programme an, wobei die meisten von kommerziellen Anbietern bereit gestellt, lizenziert und natürlich regelmässig ergänzt werden müssen.

Bei der strukturellen Analyse können zwei Verfahren unterschieden werden:

- *Formale Analyse*: Die E-Mail wird hinsichtlich des Aufbaus und Ausgestaltung des E-Header untersucht und dies mit dem vorliegenden Mitteilungsinhalt — dem MIME-Content — korreliert. Die Erfahrung sagt, dass die meisten Spam-Erzeuger versuchen, die E-Mail wie eine "übliche" — z.B. von Outlook generierte — Nachricht aussehen zu lassen, aber dann doch Fehler einbauen.  
Als weitere Kriterien lassen sich z.B. die verwendeten X-HTML-Attributen nutzen. Spam E-Mails wollen in der Regel so auffällig wie möglich sein. Ferner kann auch eine eingebettete URL auf eine Spam E-Mail deuten.
- *Inhaltliche Analyse*: Die inhaltliche Analyse besteht darin, die Bedeutung der verwendeten Worte zu erfassen (nachdem sie von eventuellen (X-HTML-) Datenmüll gesäubert wurden und gegen die gespeicherte typische Spam-Terminologie abzugleichen. Damit Worte wie "loan", "Viagra", "Erektion" nicht auch in E-Mails der damit beschäftigten Berufsgruppen als Spam identifiziert werden, bieten die meisten Systeme die Möglichkeit, spezielle Negativ- oder Positivlisten von Begriffen zu führen. Zudem wird in der Regel eine sog. Bayesean-Bewertung vorgenommen. Hierbei wird das Verhältnis von gefundenen Spam-Begriffen zu den unverdächtigen Begriffen gebildet und hiermit ein Rating durchgeführt.

Die Aufgabe der inhaltlichen Analyse ist daher sehr komplex. Ein entsprechendes System muss zum einem "Language-Aware" sein, als auch dafür Sorge tragen, dass sich die Entscheidungen nicht gegenseitig nihilieren. Hierfür gibt es Konzepte in Form sog. Entscheidungsbäume, die sicherlich in der nächsten Generation der Anti-Spam-Tools Einzug halten.

Ein weiteres Problem stellt sich, wenn in Zukunft die Aussagen der Spam E-Mails nicht in Form Text sondern z.B. einfach als eingebettete Graphiken erfolgt: Wer wollte einen Spam-Versender/Erzeuger daran hindern? Solange E-Mail Clients wie Outlook für ein eingebettetes Objekt einfach den dazu gehörigen Viewer öffnen, ist reichlich Spielraum für effektive Varianten geboten.

Alle **Post-Receipt-Spam** Verfahren haben folgende Nachteile:

- Eine komplexe und CPU-intensive Verarbeitung der empfangenen E-Mail ist notwendig.
- Die Verarbeitungskriterien müssen laufend dem Know-How der Spammer angepasst werden.
- Der Empfänger muss trotz alle dem gelegentlich in sein "Junk-Mail" Verzeichnis schauen, ob sich nicht etwa doch False-Positives hierin befinden.
- Der Spammer kann sich die (virtuellen) Hände reiben: Seine Spam E-Mail ist

zunächst vom Zielsystem entgegengenommen worden; erstes Etappenziel erreicht.

Die Vorteile dieser Verfahren besteht allerdings darin, dass

- der Empfänger entscheiden kann, ob eine E-Mail Spam ist oder nicht,
- er in der Regel selbst Einfluss auf die Entscheidungskriterien hat,
- unter Verzicht auf zusätzliche Blockierungsmassnahmen jede E-Mail den Empfänger erreicht.

Abschliessend wollen wir uns eine weitere typische E-Mail betrachten, die vom Programm Spamnix — als Plug-In für den Windows-MUA Eudora — (<http://www.spamnix.org>) analysiert und als Spam klassifiziert wurde:

```
Delivered-To: erwin@localhost
From: "Lee Baxter" <ejxhot4kwjp3@copper.net>
To: <feh@fehcom.de>
Subject: You can order Anti-depressants, weight loss meds, and
pain relief meds online with NO PRESCRIPTION cw
Date: Sun, 15 Jun 03 22:53:18 GMT
X-Mailer: Microsoft Outlook Express 5.00.2615.200
X-MSMail-Priority: High
Content-Type: text/html;
```

**ATTENTION:**

**New Viagra Soft Tab Works In 15 Mins or Less!**

**Compounded from FDA Approved Vi(a)grajFFFFAE**

**ORDER WITH NO PRESCRIPTION DISCREETLY ONLINE**

***Do It For Her!***

**Sildenafil Citrate soft tabs are compounded by licensed pharmacists using FDA Approved Vi(a)grajFFFFAE Tablets. These tablets are ground into a fine powder and processed into a sublingual medication.\***

**[CLICK HERE FOR MORE INFO](#)**

[unsubscribe here](#)

```
kqeh ec w umlfswb xsawd py abuckwtaawmek nvs ihf
```

```
SPAM: ----- Spamnix Spam Report -----  
-----
```

```
SPAM: Spamnix identified this message as spam. This report  
shows which
```

```
SPAM: rules matched the message and how many points each rule  
contributed.
```

```
SPAM:
```

```
SPAM: Content analysis details: (12.00 hits, 5 required)
```

```
SPAM: X_MSMAIL_PRIORITY_HIGH (0.4 points) Sent with 'X-  
msmail-Priority' set to high
```

```
SPAM: X_PRIORITY_HIGH (1.9 points) Sent with 'X-  
Priority' set to high
```

```
SPAM: CLICK_BELOW_CAPS (0.5 points) BODY: Asks you to  
click below (in capital letters)
```

```
SPAM: HTML_COMMENT_SAVED_URL (1.4 points) BODY: HTML  
message is a saved web page
```

```
SPAM: HTML_FONT_COLOR_UNSAFE (0.1 points) BODY: HTML font  
color not within safe 6x6x6 palette
```

```
SPAM: HTML_60_70 (0.5 points) BODY: Message is 60%  
to 70% HTML
```

```
SPAM: HTML_FONT_COLOR_RED (0.1 points) BODY: HTML font  
color is red
```

```
SPAM: HTML_MESSAGE (0.2 points) BODY: HTML included in  
message
```

```
SPAM: HTML_LINK_CLICK_CAPS (1.1 points) BODY: HTML link text  
says "CLICK"
```

```
SPAM: HTML_FONT_BIG (0.3 points) BODY: FONT Size +2 and  
up or 3 and up
```

```
SPAM: HTML_FONT_COLOR_BLUE (0.1 points) BODY: HTML font color  
is blue
```

```
SPAM: HTML_LINK_CLICK_HERE (0.1 points) BODY: HTML link text  
says "click here"
```

```
SPAM: HTML_SHOUTING4 (0.5 points) BODY: HTML has very  
strong "shouting" markup
```

```
SPAM: REMOVE_PAGE (0.3 points) URI: URL of page called  
"remove"
```

```
SPAM: FORGED_MUA_OUTLOOK (3.9 points) Forged mail pretending  
to be from MS Outlook
```

```
SPAM: MIME_HTML_ONLY (0.1 points) Message only has  
text/html MIME parts
```

```
SPAM: MISSING_MIMEOLE (0.5 points) Message has X-MSMail-  
Priority, but no X-MimeOLE
```

```

SPAM:
SPAM: Spam level: *****
SPAM: ----- End of Spamnix Spam Report -----
-----

```

Ein beliebtes und leistungsstarkes Anti-Spam-Tool ist der frei verfügbare SpamAssassin (<http://spamassassin.org/>). SpamAssassin lässt sich in nahezu jeden MTA integrieren, bietet aber auch die Möglichkeit, als Plug-In für MUAs unter Windows zu fungieren.

SpamAssassin

SpamAssassin kann über unterschiedliche Quellen bezogen werden:

1. Als vorkonfiguriertes Paket bzw. als tar-Archiv über die SpamAssassin Home-Page.
2. Als PERL-Module mittels des Aufrufs (als *root*):

```

perl -MCPAN -e shell
o conf prerequisites_policy ask
install Mail::SpamAssassin
quit

```

3. Als Bestandteil der lokalen Unix-Distribution (rpm, deb bzw. port).

SpamAssassin weist eine hohe Abhängigkeit von der installierten PERL-Version auf. Weitere vielfältige Systemabhängigkeiten und die zusätzlich vorzunehmenden Anpassungen, machen SpamAssassin nicht gerade einfach zu installieren bzw. erfolgreich aufzusetzen. Zudem existieren diffizile Unterschiede in der Konfiguration von SpamAssassin hinsichtlich der Major- aber auch Minor-Versionen.

Das ursprünglich in PERL geschriebenen SpamAssassin (CPAN: Mail::Mail-SpamAssassin) bietet mittlerweile eine Client/Server-Architektur, bei der die E-Mail zunächst von einem (in C geschriebenen) Client entgegen genommen, dann aber von Memory-residente (PERL-)Serverprozess analysiert wird. Hierdurch erreicht man, dass nicht jedesmal die umfangreich PERL-Bibliothek pro E-Mail geladen werden muss, sondern diese quasi im Hauptspeicher verbleibt.

Aufsetzen von SpamAssassin

Unter Qmail gibt es mehrere Varianten SpamAssassin einzusetzen (vgl. <http://loghog.corc.uni.edu.ni/~jorge/spamassassin.html>):

- Als Teil des Anti-Virus-Programms Qmail-Scanner (<http://qmail-scanner.sourceforge.net/>).
- Als Ersatz des Programms `qmail-queue`. Bei Qmail lässt sich dies durch das `QMAILQUEUE` Patch von Bruce Guenther (<http://www.qcc.sk.ca/bguenther/>) sowie durch das `qmail-`

**spamc** Programm John Peacock als Bestandteil von SpamAssassin erzielen:

```
qmail-smtpd => qmail-spamc (<=> spamd) => qmail-queue
```

Hierbei ist es notwendig, im Startup-Skript von **qmail-smtpd** die Environment-Variable `QMAILQUEUE="qmail-spamc"` mit der korrekten Pfadangabe zu setzen.

- In der Qmail Default-Delivery mittels des Zusatzprogramms **maildir**. **maildir** ist Bestandteil der Programmsammlung **safecat** von Len Budney (<http://budney.homeunix.net:8080/users/budney/linux/software/safecat.html> bzw. <http://freshmeat.net/releases/124946/>). Folgendes **qmail-send run**-Skript ist hierzu in der Lage.

```
#!/bin/sh
exec env - PATH="/var/qmail/bin:$PATH" \
qmail-start '|/usr/local/bin/spamassassin -L | maildir
./Maildir/'
```

- Durch (individuelle) Einbettung in der `.qmail` Datei mittels des Zusatzprogramms **maildir**.

```
|/usr/local/bin/spamassassin -L | maildir ./Maildir/
```

In diesem Fall darf keine zusätzliche Angabe von `./Maildir/` in einer weiteren Zeile erfolgen, da die Zustellung der Nachricht nun mittels **maildir** über die Pipe erfolgt.

- Das Shell-Skript **ifspamh** stellt durch seinen Aufruf innerhalb von `.qmail` Dateien (<http://www.gbnet.net/~jrg/qmail/ifspamh/>) einen Wrapper für den **spamassassin** Aufruf dar. Allerdings sehe ich in meinen Qmail-Logs öfter folgenden Eintrag (unter Nutzung der **ksh**):

```
2003-05-19 22:59:26.426539500 delivery 1196: deferral:
/usr/local/sa/bin/ifspamh[83]:_printf:_Argument_list_to
o_long//usr/local/sa/bin/ifspamh[113]:_printf:_Argument
_list_too_long/spamc_returned_temporary_failure/
```

- Als Aufruf innerhalb von **procmail** (<http://www.procmail.org/>). Hierzu muss zunächst die (lokale) `.qmail` Datei zum Aufruf von **procmail** angepasst werden:

```
# .qmail for procmail (customize .procmailrc!)
| preline /usr/bin/procmail
```

Die Einbettung von SpamAssassin erfolgt in der obligatorischen `.procmailrc` Datei:

```
# Sempel .procmailrc
```

```

:0fw
* < 265000
| /usr/bin/spamc
:0

```

Statt des monolithischen **spamassassin** PERL-Skripts kann auch die Kombination **spamc/spamd** eingesetzt werden. Der Client **spamc** (ein C-Programm) ersetzt hierbei den **spamassassin** Aufruf und verbindet sich über das Loopback-Interface und den TCP-Port 783 mit dem Daemon-Prozess **spamd** (ein PERL-Programm) — dem eigentlichen Arbeitspferd.

spamd Daemon

Hierzu muss natürlich zunächst der Daemon gestartet sein. Dies realisiert ein einfaches **run**-Skript unter den Daemontools:

```

#!/bin/sh
exec setuidgid root /usr/bin/spamd -x -L

```

Damit **spamd** eine Netzwerk-Ressource nutzen kann, muss es unter **root** laufen (via **setuidgid**); die Option **-x** besagt, dass es ohne lokalen User auskommt und mit dem Flag **-L** wird **spamd** angewiesen, nur lokale Informationen auszuwerten; insbesondere keinen DNS-Lookup vorzunehmen.

Leider war bei meiner SpamAssassin Version 2.54 und unter PERL 5.005\_03 **spamd** durch ein kleines Bug nicht lauffähig: Der Aufruf von **mkdir** wollte aufgrund einer fehlenden Verzeichnis-Mode Angaben nicht gelingen und wurde mit einem Fehler quittiert ("Not enough arguments for mkdir at /usr/bin/spamd line 878, near "\$spam\_conf\_dir"). Abhilfe schafft (etwa ab Zeile 875 in **spamd**):

```

if ( ! -d $spam_conf_dir )
{
    if ( mkdir ( $spam_conf_dir,0700 ) )
    {
        logmsg "info: created $spam_conf_dir for
$username.";
    }
}

```

Die Installation von SpamAssassin über das tar-Archiv nutzt als Standardverzeichnis **/usr/share/spamassassin/**, die Binaries werden üblicherweise in **/usr/bin/** hinterlegt. Als zentrale Konfigurationsdatei fungiert **/etc/mails/spamassassin/local.cf**. Beim Einsatz des Daemon **spamd**, werden Änderung in **local.cf** erst nach dessen Neustart wirksam.

Konfiguration

SpamAssassin kann (wie beschrieben) Systembezogen- als auch Userbezogen eingerichtet werden. Hierzu erhält jeder Benutzer ein eigenes **~/.spamassassin** Verzeichnis zur Bevorratung der Präferenzen. Zudem

unterstützt SpamAssassin auch die unter Qmail möglichen "Virtuellen Benutzer".

Aufgrund dieser Komplexität und den Abhängigkeiten vom lokalen System (**procmail**, Qmail-Scanner), kommt der System-Administrator nicht um ein genaues Studium der SpamAssassin Dokumente (**perldoc Mail::SpamAssassin::Conf**) sowie um einiges Experimentieren herum.

SpamAssassin bietet optional die Möglichkeit auf eine verteilte Spam-Datenbank Razor ("Rasierer") zuzugreifen, die von Vipul Ved Prakash ins Leben gerufen wurde (<http://razor.sourceforge.net/>) und mittels Agenten (u.a. via eines Plug-Ins für SpamAssassin) befüllt und aktuell gehalten werden kann.

Razor

Im Gegensatz zu einer RBL werden hier "Footprints" der Spam E-Mails hinterlegt, die aufgrund statistischer Analysen der E-Mails und unter Eliminierung ihrer zufälligen Anteile erzeugt werden. Dies umgeht die Notwendigkeit, eine formale bzw. inhaltliche Analyse der E-Mails vorzunehmen, sodass auch die Spam E-Mails klassifiziert werden können, die sich "Stealth" Techniken bedienen.

### 7.2.5.3 Message Digest

Das Message Digest Verfahren dreht das SMTP-Verfahren gewissermassen um. Heisst das Standard E-Mail-Verfahren:

- Akzeptiere alles — aber versuch' dich vor den "Bad Boys" zu schützen,

lautete es nun:

- Akzeptiere nichts — ausser du weisst ganz genau, dass es für dich bestimmt ist.

Beim Digesting wird somit der Absender gezwungen — wie bei einer Mailing-Liste — über einen Digest-Agent (Autoresponder) des Adressaten seine eigene E-Mail zu bestätigen. Das Message Digest Verfahren kann daher nur für individuelle Empfänger eingesetzt werden, wobei es zwei verbreitete Alternativen gibt:

1. Der Message Digest Agent verwaltet eine *Whitelist* von SMTP-Adressen (*Sender-based Delivery Confirmation*). Nur E-Mails, die von einem registrierten Absender stammen, werden sofort durchgelassen; alle anderen Absender erst durch eine zusätzliche Bestätigungs-E-Mail in die Whitelist aufgenommen. Dieses Verfahren verfolgt z.B. der bekannte Tagged Message Digest Agent TMDA (<http://www.tmda.sourceforge.net>). Zusätzliche Anti-Spam-Filter sowie ein automatisches Aging machen den TMDA für den Anwender sehr bequem, der z.B. über ein dot-qmail Plug-In integriert werden kann.
2. Eine rigidere Strategie verfolgt z.B. das Programm **qsecretary** von Dan Bernstein, das seinen Dienst in der Qmail Mailing-Liste versieht: Hier muss

jede einzelne E-Mail vom Absender bestätigt werden (*Message-based Delivery Confirmation*). Als Response auf die Original-Mail bekommt der Absender mittels des Qmail-VERP-Mechanismus einen Cookie zugeschickt, den dieser beantworten muss. Spammer erhalten so in der Regel keine Chance; ausser sie nutzen einen schnellen Autoresponder. Sicherlich ist dies einer der Gründe, warum Dan Bernstein den Quellcode des Programms bislang nicht veröffentlicht hat. Als Alternative bietet sich das Gerit Pape geschriebene **qconfirm** an (<http://smarden.org/qconfirm>), das beide Betriebsmodi unterstützt.

So bequem und effektiv diese Werkzeuge für den Empfänger von E-Mails sind, so lästig sind sie für den Absender. Manche E-Mail hat auf diese Weise deshalb ihr Ziel nicht erreicht, weil sie einfach — aus welchem Grund auch immer — nicht bestätigt wurde. Zudem verdreifacht der Message Digest Agent das E-Mail-Aufkommen prinzipiell (zumindest in der Zahl, nicht unbedingt im Volumen); keine Freude für den sparsamen E-Mail-Versender.

Andererseits gelten für MUAs, die mit einem Message Digest Agent ausgestattet sind, beim Versand der Bestätigungs-Mail im Falle von Spam vergleichbare Überlegungen wie bei Bounces: Wenige bis keine erreichen das Ziel; eine Double-Bounce Nachricht an den Postmaster ist die Regel.

Neben der freien Tools bieten mittlerweile auch kommerzielle Anbieter Anti-Spam-Dienste an. Hierbei wird von den Firmen (z.B. Brightmail <http://www.brightmail.com/>) nicht nur an Spam-Filter angeboten, sondern auch aktive "Spam-Probes" im Internet genutzt. Empfängt ein Spam-Probe eine E-Mail und klassifiziert sie als Spam wird sie mit ihren typischen Merkmalen (über einen "Footprint") in einer zentrale Datenbank eingetragen und sie auf diese Weise klassifiziert und für Nutzer des Anti-Spam-Dienstes erkennbar und somit blockierbar gemacht.

Kommerzielle Anti-Spam-Dienste

Dieses Verfahren funktioniert daher ähnlich wie eine RBL nur findet es **Post-Receipt-Spam** statt, da jede E-Mail zunächst in Empfang genommen werden muss, bevor sie klassifiziert und getagged bzw. verworfen werden kann.

#### 7.2.5.4 Tagging

Als Tagging die Hinzufügung von Informationen in der E-Mail Verstanden, die es ermöglicht, die E-Mail als Spam zu identifizieren. In der Regel werden die Merkmale entweder

- als eigenständige E-Mail Headerzeile (X-Spam-Info:),
- (im Spam-Fall) im als Teil des "Subject:"-Angabe oder aber
- (in Spam-Fall) in den Message-Body

hineingefügt. Das heute übliche Verfahren (z.B. bei SpamAssassin) ist das

Hinzufügung einer neuen Headerzeile. Ältere Version modifizieren hingegen die "Subject:"-Zeile. Hintergrund hierfür ist die Tatsache, dass nicht alle E-Mail-Clients Filter auf beliebige Headerzeilen setzen können, sodass dann das Tagging ins Leere läuft.

Problematisch beim Tagging ist, dass alle Spam-Informationen konsistent zusammengefügt werden müssen. Hierzu zählen z.B. Angaben, ob die Absender-IP auf einer Blocking List steht und per RBL gefunden wurde, welchen Stellenwert die ausgewertete SMTP-Dialoginformation (Return-Path, Hello) genießt, inwieweit und in welchem Umfang die Ergebnisse der strukturellen und inhaltlichen Analyse der Nachricht eingeflossen sind.

Vorteilhaft beim Tagging ist, dass die potentiellen Spam-Informationen zentral gesammelt und verarbeitet werden können; jeder Anwender erhält somit die eine Spam-Analyse auf gleichem Niveau. Aufgabe des Mail User Agents ist es nun, mittels geeigneter Filterkriterien das Einsortieren der Spam E-Mails vorzunehmen.

Im Daemontools `run`-Skript des SMTP-Servers können sich Aufrufe von `tcpserver`, `rblsmtpd` und natürlich `qmail-smtpd` befinden. Dan Bernstein hat seine Programme so konstruiert, dass sie über Argumente aufgerufen werden (Call Interface): `exec` ruft `tcpserver`, `tcpserver` ruft `rblsmtpd`, `rblsmtpd` ruft `qmail-smtpd` und (falls entsprechend per Patch verfügbar), `qmail-smtpd` ruft ein PAM-Modul zur SMTP-Authentisierung auf. Nach dem Start des `run`-Skriptes (oder ggf. des Run-Level-Skriptes unter `/etc/init.d/`) laufen alle Programme im selben Environment. Der Datenaustausch zwischen den Programmen kann somit effizient über Environment-Variablen erfolgen. Eine von diesen ist z.B. `$RBLSMTPD`.

Environment-Variablen

Alle zum Tagging notwendigen Informationen müssen in geeignete Environment-Variablen gesteckt werden. Anders als bei z.B. UCSPI gibt es aber für Spam keine Konventionen und keine Vorgaben. Es liegt am Anwender, alle Programme soweit zu patchen, damit sie diese Informationen einerseits bereitstellen und andererseits verwerten können. Als letztes Glied der Kette schreibt `qmail-smtpd` einen qualifizierten E-Mail Header für die einlaufende E-Mail. An dieser Stelle (`received.c`) ist die Spam-Information einzufügen.

#### 7.2.5.5 Zusammenstellung der Methoden

Die bislang vorgenommenen Überlegungen hinsichtlich der Blockade, des Filterns und des Taggings, sowie des Digesting von E-Mails sollen nun systematisch zusammengestellt werden.

Wie wir gesehen haben, gibt es

1. *Pre-Initial*-,
2. *Post-Initial*- sowie

### 3. Post-Receipt-Verfahren.

Pre- und Post-Initial-Verfahren blockieren Spam E-Mails, während alle Post-Receipt-Verfahren lediglich zum Filtern, Tagging bzw. Digesting von E-Mails nützlich sind. Aus der Sicht des Systemadministrators sind also die ersten beiden Verfahren zu bevorzugen, da sie sowohl die Last seiner Systeme als auch das Spam-Volumen im Internet reduzieren. Für den Anwender spielt dies nur eine untergeordnete Rolle. Er ist an eine an einer möglichst vollständigen Spam-Unterdrückung interessiert, die letztlich immer nur durch die Kombination mehrerer Verfahren zu realisieren ist — auch wenn diese CPU-Leistung kosten.

Für die Unterdrückung von Spam E-Mails sind zwei Grössen relevant:

- Wie *effektiv* ist eine Methode? Effektiv heisst, welchen Anteil am Spam-Aufkommen blockiert bzw. gefiltert werden kann.
- Wie *effizient* ist die Methode? Die Effizienz drückt sich durch den Anteil der False Positives sowie False Negatives aus.

Zudem haben alle Verfahren Auswirkungen auf die beteiligten Parteien:

- Den *Absender/Übermittler* der E-Mail: Bekommt er eine qualifizierte Antwort (in Falle der *False Positives*)?
- Dem *Empfänger* der E-Mail: Was hat er beim Empfang (von identifizierter Spam) E-Mail zu beachten? Besitzt er Möglichkeiten, die Klassifikation der E-Mails zu beeinflussen (Trigger)?
- Dem *Systemadministrator*? Welchen Aufwand hat er bei der Konfiguration und Administration der Anti-Spam-Verfahren?

Zwei weitere Fragen müssen bei den Verfahren gestellt werden hinsichtlich

- des *Ressourcen* (CPU, Netz, Speicherplatz) und der hiermit möglichen Denial-of-Service (DoS) Gefahren,
- der resultierenden *Bandbreite* des Spam-Aufkommens (einschliesslich Bounces und Double-Bounces).

| <i>Verfahren</i> | <b>Pre-Initial</b> |              | <b>Post-Initial</b> |                          | <b>Post-Receipt</b> |                 |                 |
|------------------|--------------------|--------------|---------------------|--------------------------|---------------------|-----------------|-----------------|
| <i>Schicht</i>   | IP                 | SMTP         | IP                  | SMTP                     | Anwendung (Queue)   | Anwendung (MDA) | Anwendung (MUA) |
| <i>Name</i>      | Remote RBL         | SPAM-CONTROL | IP-Ausschlussliste  | Absender-Ausschlussliste | Filter              | Filter          | Filter          |

| <b>Verfahren</b>                  | <b>Pre-Initial</b>    |                                | <b>Post-Initial</b> |                      | <b>Post-Receipt</b> |               |             |
|-----------------------------------|-----------------------|--------------------------------|---------------------|----------------------|---------------------|---------------|-------------|
| <i>Beispiel</i>                   | RBL-SMTPD             | badmail-from/<br>badhelo       | Spamtrap            | RTS                  | procmal             | Spam-Assassin | TMDA        |
| <i>Trigger</i>                    | extern                | intern,<br>Konfigurationsdatei | self                | self/Em-pfänger      | extern/-Em-pfänger  |               | Absender    |
| <i>Risiko (False Negatives)</i>   | mittel                | gering                         | sehr gering         | sehr gering          | gross               | gering        | gross       |
| <i>Effektivität</i>               | 30%                   | 10%-30%                        | 50%                 | 70%                  | 50% --              | 90%           | 100%        |
| <i>Effizienz</i>                  | 100%                  | 100%                           | 90%                 | 90%                  | 50% --              | 90%           | 100%        |
| <i>Empfänger-Aktion</i>           | keine                 | keine                          | möglich             | möglich              | möglich             | möglich       | möglich     |
| <i>Absender (False Positives)</i> | keine bzw. URL        | SMTP Protokollfehler           | URL                 | SMTP-Protokollfehler | keine               | keine         | Return-Mail |
| <i>Administration</i>             | gering                | mittel                         | gering              | gering               | gross               | mittel        | Empfänger   |
| <i>Bandbreite</i>                 | gering/<br>DNS-Lookup | sehr gering                    | gering              | gering               | hoch                | hoch          | sehr gross  |

Tabelle 7.2-2: Zusammenstellung der Blockade- und Filtermassnahmen für Spam E-Mails.

## 7.2.6 Zusätzliche Anforderungen

### 7.2.6.1 Anforderung an den E-Mail-Provider

Egal ob blockiert oder gefiltert, für den E-Mail Provider stellt sich die Anforderung, nicht nur die durchgelassenen E-Mails zu protokollieren, sondern auch im besonderen Masse auch die blockierten bzw. gefilterten. Hierzu sagt RFC 2505 in §2.4:

*"The MTA SHOULD be able to log all anti-relay/anti-spam actions. The*

*log entries SHOULD contain at least:*

- o Time information.*
- o Refusal information, i.e. why the request was refused ("Mail From", "Relaying Denied", "Spam User", "Spam Host", etc).*
- o "RCPT To:" addresses (domains).  
(If the connection was disallowed at an earlier stage, e.g. by checking the SMTP\_Caller IP address, the "RCPT To:" address is unknown and therefore cannot be logged).*
- o Offending host's IP address.*
- o Offending host's FQDN hostname.*
- o Other relevant information (e.g. given during the SMTP dialogue, before we decided to refuse the request)."*

Zur Kontrolle und Verständnis des Spam-Niveaus muss der Systemadministrator des MTA eine qualifizierte Analyse seiner Logfiles vornehmen. Mittels meines SPAMCONTROL-Patches für Qmail sowie dem Zusatzprogramm **newanalyse** lässt sich dies im Falle von Qmail sehr bequem realisieren. Voraussetzung ist allerdings, dass die Log-Aufzeichnung mittels **multilog** von Dan Berstein erfolgt.

Zusätzlich sind die Logfiles — zumindest für einen bestimmten Zeitraum — zu archivieren. Auch hier bietet **newanalyse** ein einfaches Verfahren an. Im Gegensatz hierzu rotieren zwar auch die per **syslogd** generierten Logfiles (in der Regel `/var/log/maillog`), doch werden üblicherweise nur die letzten 10 Tage aufgehoben. Hier muss der Systemadministrator händisch für die Archivierung der Loginformation sorgen.

#### **7.2.6.2 Anforderungen an eine künftige E-Mail Infrastruktur**

Es gibt Pessimisten, die annehmen, dass SMTP-E-Mail in einigen Jahren aufgrund des Spam zusammenbrechen wird. Dies gilt unter der Voraussetzung, dass der Spam-Anteil stetig steigen wird. Natürlich kann dies kein Mensch voraussagen. Andererseits steigen sowohl die Anzahl der E-Mail-Teilnehmer, als auch die Kapazitäten zur Verarbeitung der E-Mails (Bandbreite, Leistung der MTAs etc.). Inwieweit sich die Faktoren kompensieren oder auch verstärken, kann nicht qualifiziert abgeschätzt werden. Fraglich ist auch, ob es nicht aufgrund des zu erwartenden "Sättigungsverhalten" hinsichtlich von Spam dazu kommt, dass die Effektivität für die Spammer so sehr fällt, dass das Geschäft nicht mehr lukrativ ist.

Andererseits werden Stimmen lauter, dass eine neue E-Mail-Infrastruktur geschaffen werden müsste, die weniger Spam-Anfällig ist. Drei Ansätze wurden bislang gemacht:

1. Abkehr von der SMTP-E-Mail; hierzu gab/gibt es eine Initiative von Dan Bernstein, "Internet Mail 2000" IM2000 (<http://cr.yp.to/im2000.html>). Hierbei reden wir immer von drei Bausteinen: 1. Dem E-Mail-Protokoll selbst, 2. dem Aufbau der E-Mail und 3. den Zulieferprotokollen à la POP3/IMAP4.
2. Ergänzung des (E-)SMTP-Protokolls; als Beispiel hierfür fungiert TORO: "Trust of Reliable Origin" von Marc-Andre Pelletier (<http://www.ietf.org/internet-drafts/draft-smtp-trust-00.txt>).
3. Ergänzung des DNS-Protokolls zum Reverse-Lookup der IP-Adressen von MTAs, was als RMX oder manchmal auch als XM bezeichnet wird; hierzu gibt es einen Draft von Hadmut Danisch (<http://www.ietf.org/internet-drafts/draft-danisch-dns-rr-smtp-02.txt>).

Während Dan Bernstein's Vorstoss bislang weitgehend unbeachtet blieb, besitzen die letzteren beiden einen IETF Draft-Status. Kritisch muss an dieser Stelle beleuchtet werden:

- Wir haben kein eigentliches Problem des SMTP-Protokolls hinsichtlich Spam. SMTP funktioniert so gut, wie es geplant ist: Als Übertragungs- und Nachrichten-Aufbau-Protokoll. Die aktuellen Versionen (RFC 2821/2822) vom Autor Klensin haben zum Verständnis und Weiterentwicklung nur wenig beigetragen.
- Ergänzungen, wie die SMTP-Authentication, haben relativ wenig gebracht, weil Spam kein Problem der *Authentisierung*, sondern eins der *Authorisierung* ist: Der SMTP-Sender kann in der Regel senden *was* und *wohin* er will; jeder SMTP-Empfänger muss akzeptieren, was er bekommt (nur nicht-authorisiertes Weiterleiten darf unterbunden werden). Die Situation ist also recht disparitatisch.
- Der Vorschlag von M.A. Pelletier löst vielleicht die Frage der Authentisierung über Relays; nicht jedoch die der Authorisierung. Ferner vermischt der Draft die bislang beim SMTP-Protokoll sorgfältig vorgenommene Trennung zwischen Protokollelementen zur Übertragung (RFC 821) und solche zur Beschreibung des Nachrichten-Formats (insbesondere des E-Mail-Headers; RFC 822).
- Die DNS-Ergänzung von H. Danish ist demgegenüber sogar noch kontraproduktiv: Sie bricht das Forwarding von E-Mails mit dem Original-SMTP-Abender über MTAs und sie erzeugt unerwünscht grosse DNS-Pakete; ferner ist sie für ISPs nur schlecht administrierbar, weil diese ständig die IP-Adressen ihrer MTAs im DNS aktualisieren müssen.

Die Authorisierung von Sendern lässt sich nicht über technische, sondern nur über administrative Massnahmen regeln. Ein Vorschlag besteht darin, für jede Domain mit einem SMTP-Sender, die möglichen MTAs mit ihrer jeweiligen Reverse-IP-Adresse über einen trivialen DNS TXT-Record kenntlich zu machen:

```
102.23.4.195.in-addr.arpa TXT MTA=yes
```

Hierüber kann sofort (Pre-Receipt) mittels der IP-Adresse überprüft werden, ob ein SMTP-Sender per DNS authorisiert ist. Dies ist unabhängig davon, *welche* E-Mail er verschickt und *welchen* SMTP-Absender diese besitzt. Hierfür müsste der SMTP-Server so modifiziert werden, dass per Default ein DNS TXT Lookup mit der Reverse IP-Adresse des sendenden SMTP-Clients vorgenommen wird. Bei **qmail-smtpd** kann für Clients, bei denen die Environment-Variable `$RELAYCLIENT` gesetzt ist, auf diesen Lookup verzichtet werden. Für E-Mail-Sender, die per Dial-Up (d.h. ständig wechselnden IP-Adress) über den MTA ihrer Firma senden möchten (*Roaming User*), bietet sich an, SMTP-Authentication zu verwenden. Das Verfahren ist weiterhin kompatibel mit dem beliebten *POP-before-SMTP*-Verfahren.

Die Tatsache, dass Protokolle wie POP3 und IMAP4 nur zum Abholen der E-Mail, nicht aber zu Verschicken geeignet sind, sondern letzteres über das SMTP-Protokoll erfolgt, ist sicherlich einer der Mitverursacher der heutigen Spam-Misere.

### 7.2.6.3 Organisationen und nützliche Links

Zum Abschluss möchte ich noch auf einige mehr-oder-weniger nützliche Quellen im Internet hinweisen, hinter denen sich Organisationen befinden, die sich der Spam-Abwehr bemühen. Diese Liste ist bei weitem nicht vollständig:

- Die "offizielle" Anti-Spam Research Group (ASRG) des IETF: <http://www.irtf.org/charters/asrg.html>
- Das Internet Mail Consortium (imc): <http://www.imc.org/>
- Die deutsche Anti-Spam Webseite von Florian Klein (alias "DocSnyder"): <http://www.antispam.de/>
- <http://www.spamfaq.net/> (nicht mehr Online).
- Der Anti-Spam-Dienst SpamCop: <http://www.spamcop.net/>
- Der Anti-Spam-Dienst SPAMHAUS: <http://spamhaus.org/>
- Die Webseite des Mail Abuse Prevention System LLC (MAPSSM): <http://mail-abuse.org/>

Weitere interessante URLs finden sich unter dem jeweiligen Thema in diesem Abschnitt; inwieweit diese aber Bestand haben, ist natürlich ungewiss.