

e-Mail-Postfächer zu 99% spamfrei? Strategien, Methoden, Wirksamkeit

7. Juni 2004

Rechtlicher Hinweis

Dieser Beitrag ist lizenziert unter der UVM Lizenz für die freie Nutzung unveränderter Inhalte.

Zusammenfassung

Alle stöhnen, alle leiden. Alle? Wirklich alle? Manche Provider nennen Ihre Schutzmaßnahmen schon SSPAM-Schutz", wenn sie nur 30% der unerwünschten Mails herausfiltern, andere Provider filtern gleich dutzendweise normale legale Mails in die Spamverdachtsordner. Dabei ist es eigentlich recht einfach. SPAM-freie Postfächer sind keine Utopie, sondern bei guten Providern selbstverständlicher Standard. Wir zeigen am Beispiel des Providers "JPBerlin.de", wie man mit einer geschickten Kombination von Überprüfungen, Filtersystemen und einem bißchen "künstlicher Intelligenz" mittels Postfix, amavisd-new und anderer Ansatzmöglichkeiten einen sicheren Mailserver konzipiert, der zu >99% SPAM-frei ist, kaum fehlerhaft filtert und nebenbei so virensicher ist, wie nur möglich.

Zum Nachbau empfohlen.

Der Vortrag wird von Peer Heinlein gehalten, der mit dem Buch "Das Postfix-Buch" (früher SuSE Press, jetzt 2. Auflage Open Source Press) als erster und bislang einziger ein umfassendes Buch zum Thema Sichere Mailserver mit Linux" geschrieben hat. Hinweis: Der Vortrag tritt mit dem Nachfolgenden Vortrag als Doppelvortrag auf; beide Vorträge bauen aufeinander auf.

1 Ursachen und Quellen von Spam

Um Spam zu bekämpfen und zuverlässig abzublocken, lohnt es sich zuerst zu analysieren, woher Spam heutzutage kommt und wo die Ursachen liegen:

1.1 „Open Relays“

- Schlecht konfigurierte Server nehmen Mails aus dem Netz an,
- obwohl sie an externe Adressen gehen (Öpen Relay")
- Oft genug noch default-Einstellung...

1.2 Dialup-User

- Bandbreite auch als dialup mittlerweile überall zu haben

1.3 Professionelle Spam-Hoster

- Außerhalb unserer Gerichtsbarkeit

1.4 Web-Proxy-Server, cgi-Mail-Scripte

- Methode "CONNECT" erlaubt TCP/IP-Connect an lokalen Mailserver, der erlaubt Mail, weil ja "lokaler Client" aus dem LAN
- cgi-Scripte, die e-Mails erlauben, oft ohne Schutz (formmail.pl)

1.5 Viren mit SMTP-Engine

- Neueste Entwicklung
- Unvorstellbar große Anzahl und Kapazitäten (siehe auch "distributed computing/SETI")

2 Lösungsansätze

Aus den Ursachen und Quellen des Spams kann man nun eine Strategie ableiten, die möglichst jede einzelne Ursache aufgreift und mit ihr fertig werden sollte. Im normalen Mailserverbetrieb ohne technische Tricks :-)

stehen folgende Ansatzmöglichkeiten zur Verfügung:

- Bekannte Spammer-IPs blocken (RBL)
- Dialup-IP-Bereiche blocken (DUL)
- Access-Listen nach Absenderadressen
- Body-/Headerfilter
- Spam nach Merkmalen erkennen (heuristisch)
- Verteilte "human detection" Systeme (razor)

Schauen wir uns die einzelnen Bereiche näher an.

3 Realtime Blacklists / Dialup User List

- Spezielles "Reverse Lookup" blockt IP-Nummern über DNS-Abfragen
- RBL nicht ganz unproblematisch...
 - P: Unbeständige RBL-Listenbetreiber
 - P: Angriffe der Spammer gegen RBL-Listen
 - P: DoS möglich (Osirusoft)
 - P: Geht nicht mit fetchmail
- ...aber sehr wirkungsvoll und darum absolut sinnvoll.

4 Body- und Headerchecks

Body- und Headerchecks filtern nach einzelnen Wörtern (zu unsicher!), Textpassagen (passabel) oder über Regular-Expression-Ausdrücke auch gegen bestimmte Pattern (z.B. Attachments mit doppelter Dateiendung). Diese Checks sind eigentlich umständlich und unflexibel, da Spam-Mails einzeln aufgenommen werden müssen, was eine permanente Pflege bedeutet. Sie helfen aber ungemein gegen wiederkehrende penetrante Spammer, die es schaffen alle anderen Schutzmaßnahmen zu unterlaufen und die wir so manuell blocken können.

4.1 Beispiele für Bodychecks

```
/Willkommen beim Lucky7Casino/ REJECT  
/schlechte Schufa-Auskunft? Bonitätsprobleme? Dann wählen Sie doch/ REJECT  
/The Weekend Pill - Xialis is safer, quicker, lasts longers/ reject
```

4.2 Beispiele für Headerchecks

```
/^Subject:*Weihnachtsbaeume ab EURO.* / REJECT  
/^subject: =?big5?Q/ REJECT  
/^X-AD2000-(Serial|Register):/ REJECT  
/^X-Mailer:.*\b(Aristotle|Avalanche|Blaster|Bomber|DejaVu|eMerge|Extractor|UltraMail)\b/ REJECT
```

5 Heuristische Methoden zur Spam-Erkennung

Am flexibelsten und erfolgversprechensten ist die Beurteilung einzelner Mails nach einer Summe von Merkmalen, die für oder gegen Spam sprechen. Aus ihnen läßt sich die Wahrscheinlichkeit errechnen, ob die geprüfte e-Mail normal oder Spam ist.

Anders als bei Body- und Headerchecks läßt sich so auch Spam erkennen, der zum ersten mail auftritt und darum in manuellen Filtern noch nicht aufgenommen sein kann.

Für die Beurteilung der Mails stehen uns viele, viele Kriterien zur Verfügung:

- Bestimmte Wörter / Blacklists
Die üblichen Begriffe, Kombinationen besonders gewertet.
- Verdächtige URLs im Text
clickme, removeme, unsubscribe
Getarnte Hostnamen, codierte IP-Adressen
- Schlechte Fakes – auch Spammer machen Fehler
Falsch gefälschte Mailagents, Zeitzonen, Received-Zeilen, kaputte MIME-Codierung im Mailbody
- Weitere Checks
RBL / DUL-Check
Massive Häufung von HTML-Tags (Schriftfarbe, Schriftgröße)
angehängter Zahlen- / Buchstabencode in Betreff und Mailabspann
Ungültige HTML-Tags um Spam-Wörter zu zerhacken
base64-Codierung
- Lernender Filter möglich!
Zwei IMAP-Ordner, SSpam und "Ham", Lernprozeß bildet "fuzzy checksum" auch Abweichungen zu erkennen.
- Perfekt: SpamAssassin und Razor
P: Spammer beginnen ihre Mails gegen SA abzugleichen!

Die Beurteilung einer solchen Mail kostet natürlich Rechenzeit und ist aufwendig, darum versuchen wir später möglichst wenige e-Mails bis zur Prüfung durch SpamAssassin durchzulassen.

6 Erwischt - Beispiel Nummer 1

Dieses Beispiel wurde völlig zweifelsfrei als Spam erkannt (Score dieser Mail: 13.0, Kill-Score: 5.1...). Auch wenn der Text der Nachricht base64-codiert ist (und darum Body- und Headerfilter nicht greifen konnten!) hat SpamAssassin diese Mail decodiert und geprüft - die Auflistung zeigt, welche verdächtigen Prüfungen auf diese e-Mail alles zutrafen: Falsches Maildatum, außerordentlich viel HTML, große Schriften, kein HTML-Title, keine ASCII-Version und ein paar „böse Wörter“ ...

```
Return-Path:
Delivered-To: spam-quarantine
X-Envelope-To:
X-Envelope-From:
X-Quarantine-id:

Received: from chemeng.chmt.wits.ac.za (cable255a124.usuarios.retecal.es
[212.183.255.124])
    by brainy.jpberlin.de (Postfix) with ESMTP id 1BD1CD71C9
    for ; Mon, 19 Jan 2004 21:37:40 +0100 (CET)
Message-ID:
From: "Rudy Barron"
To: kusano@jpberlin.de
Subject: feeling down about the slze of your johanson...
Date: Wed, 21 Jan 2004 20:37:08 +0000
MIME-Version: 1.0
Content-Type: text/html
Content-Transfer-Encoding: base64
X-Spam-Status: Yes, hits=13.0 tag1=3.0 tag2=5.1 kill=5.1 tests=BAYES_99,
  DATE_IN_FUTURE_24_48, HTML_70_80, HTML_FONTCOLOR_BLUE, HTML_FONTCOLOR_RED,
  HTML_FONT_BIG, HTML_MESSAGE, HTML_TITLE_UNTITLED, IMPOTENCE,
  MIME_BASE64_TEXT, MIME_HTML_NO_CHARSET, MIME_HTML_ONLY, PENIS_ENLARGE
X-Spam-Level: *****
```

```
PCFET0NUWVBFIEhUTUwgUFVCTEldICItLy9XM0MvL0RURCBIVE1MIDQuMDEg
VHJhbnNpdGlvbmFsLy9FTiI+DQo8aHRtbD4NCjxoZWZkPg0KPHRpdGx1PlVu
dG10bGVkIERvY3VtZW50PC90aXRszT4NCjxtZXRhIGh0dHAtZXF1aXY9IkNv
bnR1bnQtVHlwZSIgY29udGVudD0idGV4dC9odGlsOyBjaGFyc2V0PWlzby04
ODU5LTEiPg0KPC9oZWZkPg0KDQo8Ym9keT4NCjxwIGFsaWduPSJsZWZ0Ij48
c3Ryb25nPjxmb250IGNvbG9yPSIjMDAzM0NDIiBzaXplPSI0Ij5XYTxrY3Vm
cmZxMWVkbmV0Mj5udD48Zm9udCBzaXplPSI0Ij4gDQogIEEgPGZv
bnQgY29sb3I9IiNNGRjAwMDAiPkJpZzxrNmYwdTliZDZqeDhjam0+PGtkN28z
YWkxM3J2aTJnMj5nZXI8L2ZvbnQ+IFB1PGt0cXUwaJZxcjV1YjM2Pm5pPGth
dmM2cTEExaXl1PnM/PC9mb250Pjwvc3Ryb25nPjwvcD4NCjxwIGFsaWduPSJs
```

7 Erwischt - Beispiel Nummer 2:

Das zweite Beispiel zeigt eine in den letzten Monaten neue aufgekommene Masche der Spammer: Sie garnieren den Text zufällig mit ungültigen schließenden HTML-Tags. Diese werden in der Darstellung einfach ausgeblendet, so daß das davor und danach stehende Wort zusammengesetzt sichtbar wird. Textfilter, die nach diesen Wörtern („Viagra“) suchen, laufen also ins Leere. Zudem varrieren damit die einzelnen Mails sehr stark, was die Überprüfung durch eine „fuzzy checksum“ erschwert oder unmöglich macht.

Darüber hinaus hat SpamAssassin noch eine ganze Reihe weiterer Spam-Merkmale entdeckt.

Return-Path: <02gfxfi@qsl.net>
Delivered-To: spam-quarantine
X-Envelope-To: <eine-welt-infoladen@jpberlin.de>
X-Envelope-From: <02gfxfi@qsl.net>
X-Quarantine-id:
<spam-00542ec198f4edf7dbbb2548c09ee5ab-20040109-095117-17804-05>
Received: from 62.8.206.147 (unknown [211.41.251.30])
by brainy.jpberlin.de (Postfix) with SMTP id 22099D60B3
for <eine-welt-infoladen@ewil-koepenick.de>; Fri, 9 Jan 2004
09:51:02 +0100 (CET)
Received: from [40.17.5.212]
by 62.8.206.147;
Fri, 09 Jan 2004 15:40:34 -0400
Message-ID: <8u3tfz6310\$hku\$\$6y657-0@06b.a.8tn>
From: "Kelly Bright" <02gfxfi@qsl.net>
Reply-To: "Kelly Bright" <02gfxfi@qsl.net>
To: eine-welt-infoladen@ewil-koepenick.de
Subject: Fwd: V|@gra, Vali(u)m, X(a)n@x Diet Pills Here. 95534293939 azg
Date: Fri, 09 Jan 04 15:40:34 GMT
X-Mailer: ELM [version 2.5 PL3]
MIME-Version: 1.0
Content-Type: multipart/alternative;
boundary="._AAF7_.3.02_5DAF9E"
X-Priority: 1
X-MSMail-Priority: High
X-Spam-Status: Yes, hits=17.7 tag1=3.0 tag2=6.3 kill=6.3 tests=BAYES_90,
DATE_IN_PAST_03_06, DATE_SPAMWARE_Y2K, FORGED_RCVD_NET_HELO, HTML_60_70,
HTML_MESSAGE, HTTP_ESCAPED_HOST, HTTP_EXCESSIVE_ESCAPES, MIME_HTML_ONLY,
MIME_HTML_ONLY_MULTI, MISSING_MIMEOLE, MISSING_OUTLOOK_NAME,
X_MSMail_PRIORITY_HIGH, X_PRIORITY_HIGH
X-Spam-Level: *****

--._AAFx7_.3.02_5DAF9E
Contentx-Type: text/html;
Contentx-Transfer-Encoding: quoted-printable

/strong>Imp</embarrass>rov</abelian>ing t</clarence>he qua</=
blackstone>li</cyclist>t</age>y of peo</criteria>ple's=
liv</affirmative>es is wh</lilac>at pr</katmandu>escr</=
athena>iption medic</laboratory>at</sidecar>ions ar</=

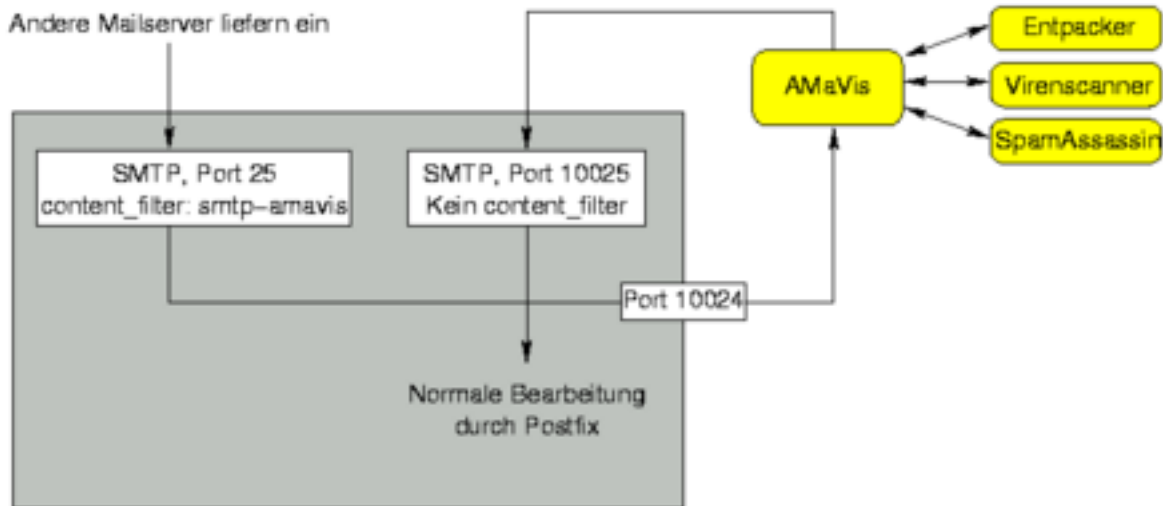
8 Implementierung in einem MTA (Postfix)

Viren- und Spamkiller bringen heute meist eine eigene kleine SMTP-Engine mit, so daß sie sich recht einfach in beliebige bestehende Mailserver integrieren lassen. Einzige Voraussetzung ist, daß man dem Mailserver beibringen können muß, alle e-Mails an einen „Relayhost“ weiterzuleiten. Das kann ein völlig eigenständiger Rechner sein, bei dem AmaVis & Co auf Port 25 lauschen, das kann aber auch „localhost“ sein und dort oft auf Port 10024.

Eine geprüfte und für okay befundene e-Mail kann dann wahlweise an einen zweiten Mailserver weitergereicht werden - sozusagen als eine Kette - oder aber beim ursprünglichen Mailserver auf einem eigens

dafür reservierten Port („localhost:10025) wieder eingespielt werden, damit dieser die e-Mail in die Postfächer zustellt.

AmaViS ist eigentlich als Virenkiller gestartet, er zerlegt e-Mail und Attachements in alle Einzelteile und gibt sie einem normalen Desktop-Virenkiller zur Prüfung. AmaViS integriert aber auch wunderbar und sehr performant SpamAssassin, so daß amavisd-new heute erste Wahl für die Einbindung von SpamAssassin sein sollte - virengeprüfte e-Mails sind dabei ein nettes, willkommenes Abfallprodukt.



(Grafik

steht unter Copyright - mit freundlicher Genehmigung entnommen aus „Das Postfix-Buch“, Open Source Press)

9 Statistik aus der Praxis

Ich möchte Ihnen eine kleine Statistik eines unserer Mailserver vorstellen, bei dem wir beispielhaft den Februar 2004 ausgewertet haben um zu sehen, welche Methode wieviele Spam-Mails blockte.

Von 2.576.257 Mails wurden 1.040.492 Mails als Spam enttarnt und geblockt (40%). Von diesen 1.040.492 Spams wurden entfielen auf die einzelnen Stufen:

- Stufe 1: RBL / DUL / RHSBL
728.111 Mails (ca 70 % des Spams)
- Stufe 2: Access-Listen
102.512 Mails (ca 10 %)
- Stufe 3: Body- und Header-Checks
116.316 Mails (ca 11 %)
- Stufe 4: AMaViS + SpamAssassin ("künstliche Intelligenz"...)
93.553 Mails (ca 9 %)

Das Interessante dabei ist, daß Stufe 1 und 2 vor der Übertragung der eigentlichen e-Mail ansetzen, sich also rein auf die Angaben des SMTP-Dialogs beziehen. Fällt dort die Entscheidung die e-Mail gar nicht annehmen zu wollen, wird die eigentliche e-Mail gar nicht mehr übertragen - das spart immens Traffic.

Nur Stufe 3 und 4 setzen nach der erfolgten Übertragung der e-Mail an, da wir dort ja den Inhalt der e-Mail beurteilen müssen.

Eine kluge Kombination dieser vier Stufen sorgt dafür, daß wir möglichst wenig Traffickosten haben und zugleich möglichst wenig e-Mails aufwendig durch Textfilter oder SpamAssassin prüfen müssen - das spart Ressourcen! Über 80% konnten wir bereits vorher abblocken .

10 Rechtslage Spam

Leider ist es für Privatanwender kaum möglich gegen Spammer vorzugehen, selbst wenn ihre Identität bekannt sein sollte. Rechtliche Ansatzpunkte ergeben sich defacto bislang nur aus dem Wettbewerbsrecht und stehen damit nur Konkurrenten des Spammers oder Verbraucherschutzvereinen offen.

10.1 Im privaten und gewerblichen Bereich

- Unzulässig, parallele zu Faxwerbung
- Schutz aber nur nach §§ 823 I, 826 BGB
- Kaum sinnvoller Rechtsschutz, da nur SSchadenëinklagbar
- Problem: Prozeßkosten, Dauer, Risiko Zivilrichter!

10.2 Zusätzlich im gewerblichen Bereich

- Spam-Versand ist unlauterer Wettbewerb
- Konkurrenz und Verbraucherverbände können über §§ 1,3 UWG abmahnen
- Abmahnkosten, strafbewehrte Unterlassungserklärung
- Einzig wirkungsvoller Schutz
- Problem: Anwalt verdient, man selber kriegt nix, trägt aber Risiko
- Problem: Versender im Ausland

11 Rechtliche Probleme

Der Einsatz von Spam-Schutzmaßnahmen ist für Postmaster überraschend schwierig - oder gefährlich.

11.1 § 206 StGB, Verletzung des Post- oder Fernmeldegeheimnisses

- Mitarbeiter eines Unternehmens, das geschäftsmäßig Post- oder Telekommunikationsdienste erbringt...
- ...der unbefugt eine einem solchen Unternehmen zur Übermittlung anvertraute Sendung unterdrückt...
- ...wird mit Freiheitsstrafe bis zu fünf Jahren oder mit Geldstrafe bestraft.
- (Gilt auch für Unternehmen im Umfeld oder Zulieferer.)

11.2 Betriebsrat in Firmen

- Mitbestimmungsrecht, §87 Betriebsverfassungsgesetz
- Kann zustimmen, ersetzt individuelle Vereinbarung mit AN
- Konflikt mit Admin: Wann ist allg. technischer Schutz ein individueller Eingriff?
- Betriebsrat muß gefragt werden, muß aber auch zustimmen, wenn alles zulässig ist (Chance zur Überwachung).

11.3 Lösungen

- Eindeutige schriftl. Regelung mit Betriebsrat, Arbeitnehmer oder Kunden
- "Tag only", Kunde kann selber filtern
- öpt inin das Schutz-System

12 Finanzielle Aspekte

Spam wird zu oft noch „duldsam ertragen“. Die Konfirguration eines Mailservers, der Spam wirksam herausfiltert, erfordert etwas Übung und Aufwand. Problematisch ist es dabei den Spagat hinzukriegen, daß eben nur Spam und nicht auch normale e-Mails herausgefiltert werden. Zudem fehlt im betrieblichen Streß oft genug Man-Power um diese zusätzliche Aufgabe zu bewältigen.

Schaut man sich an, welche Kosten durch Spam verursacht werden, sieht die Rechnung jedoch schnell anders aus. Die Kosten für einen externen Consultant o.ä. sind „Peanuts“ :-) gegenüber den eingesparten betrieblichen Kosten.

12.1 Kosten von Spam

- Traffic
830.000 gefilterte Mails a 20 Kbyte = ca. 16.5 Gbyte.
Gleiche Menge nochmal bei externem Abruf durch POP3/IMAP !
33 Gbyte a 3 EUR = 100 EUR im Monat, 1200 EUR im Jahr.
- Arbeitszeit Arbeitnehmer?
- Support-Aufwand Techniker?
- Schaden durch versehentlich gelöschte echte Mails?

12.2 Abschnitt

13 In eigener Sache...

Wir möchten Postmaster und andere IT-Verantwortliche dazu ermuntern, sich des Spam-Problems fachmännisch anzunehmen und Schutzmaßnahmen dagegen aufzubauen - wie man sieht erfordert es weder viel Aufwand, noch vieler Kosten um einen „Grundschutz“ zu gewährleisten. Die ersten 80% Schutz sind am einfachsten...

Natürlich stehen auch wir mit Rat & Tat zur Seite und können Mailserver professionell, routiniert und schnell remote administrieren, updaten, erweitern und prüfen. Zudem haben wir über unser eigenes Rechenzentrum in Berlin die Möglichkeit ganze Maildomains über unsere Server zu relayen von Spam zu befreien, bevor wir Sie gesäubert an den Kunden weiterreichen. Gerne stehen wir für individuelle Anfragen, Beratung und Consulting zur Verfügung:

Heinlein & Partner Linux Consulting

<http://www.heinlein-partner.de>

Von uns stammt das Buch „Das Postfix-Buch - Sichere Mailserver mit Linux“, das früher bei SuSE Press und jetzt in der 2. Auflage bei Open Source Press erschienen ist:

<http://www.postfixbuch.de>

<http://www.opensourcepress.de>

Wir führen bundesweit Schulungen über den Betrieb sicherer e-Mail-Server auf der Basis von Postfix durch:

<http://www.opensourceakademie.de>