

---

# Spam-Bekämpfung mit Open Source-Software

Jochen Topf

14.09.2005

OpenSaar 2005

# Einführung

---

- Viele Antispam-Maßnahmen
- Unüberschaubare Menge an Open Source-Programmen
- Implementierung (fast) aller Maßnahmen als OS vorhanden
- Vorstellung verschiedener Maßnahmen und Programme
- Client und Server

# Client oder Server?

---

- Client
  - Leicht einzurichten und zu konfigurieren und trainieren
  - Flexibel (passend für Anwender)
- Server
  - Spart Ressourcen
  - Mehr Filterkriterien
  - aber: Zugriff auf Mailserver erforderlich

# Open Source?

---

- Client
  - Sehr gute Programme
  - Einfache Benutzung
  - Kommerzieller Software mindestens ebenbürtig
- Server
  - viele hervorragende Bausteine
  - kaum integrierte Software

# Im Client

---

- Fast alle Mail-Clients bieten Spam-Filter an
- Funktionen zur Sortierung/Markierung (nicht nur für Spam!)
- Typischerweise
  - Einbindung externer Programme oder
  - Verwendung von Bayes-Verfahren...

# Bayes-Verfahren

---

- Statistische Inhaltsanalyse
- Basiert auf Worthäufigkeiten
- Lernfähig. Wird durch einfaches Vorlegen von Spam und Nicht-Spam-Mails trainiert
- Fortgesetztes Training erforderlich
- Sehr effektiv (>98% Erkennungsrate)
- Relativ langsam

# Mozilla Thunderbird

---



- Konfigurierbare Header-Filter
- Whitelisting
- Bayes-Filter (“Adaptive Filter”)
- Einfaches Training des Bayes-Filters
- Flexible Behandlung von Spam (verschieben, löschen, ...)



# Thunderbird Filter

Filter name:

For incoming messages that:

Match all of the following  Match any of the following

Subject	contains	[Spam]
X-Spam-Score	contains	+++++

Perform these actions:

<input checked="" type="checkbox"/> Move to folder:	SPAM on jochen@remote.org	<input type="button" value="New Folder..."/>
<input type="checkbox"/> Copy to folder:	jochen@remote.org	<input type="button" value="New Folder..."/>
<input type="checkbox"/> Label the message:	Important	
<input checked="" type="checkbox"/> Change the priority to:	Lowest	
<input checked="" type="checkbox"/> Set Junk Status to:	Junk	
<input type="checkbox"/> Mark the message as read		
<input type="checkbox"/> Flag the message		

# Thunderbird Junkmail filter

Configure Junk Settings for:

Settings | Adaptive Filter

**White Lists**

Do not mark messages as junk mail if the sender is in my address book:

**Handling**

Move incoming messages determined to be junk mail to:

- "junk" folder on:
- Other:

Automatically delete junk messages older than  days from this folder

When I manually mark messages as junk:

- Move them to the "junk" folder
- Delete them

When displaying HTML messages marked as junk, sanitize the HTML

**Logging**

View and configure junk mail logging.

# KDE KMail

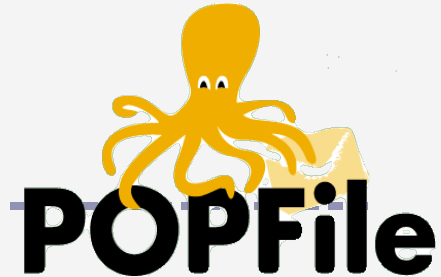
---



- Erlaubt Filter ähnlich wie bei Thunderbird, aber noch flexibler
- Normale Filter und Download-Filter, die vor dem Download greifen
- Keine eingebaute Antispam-Funktion
- Kann externe Antispam-Programme wie SpamAssassin und Bogofilter nutzen
- (ähnlich auch Evolution)

# POPFile

---



- Nutzung, wenn man weder Client noch Server verändern kann
- Arbeitet als POP-Proxy
- Funktioniert mit jedem Client
- Bayes-Filter
- Konfiguration und Training über Webseiten
- Windows und Cross-Plattform Version

# Auf dem Server

---

# MTA (Mail Transfer Agent)

---

- Der MTA ist für das Routing und das Zustellen von E-Mails zuständig
- Idealer Ort für die Spam- und Virenfilterung
- Filterung noch während der SMTP-Verbindung möglich
  - Ablehnen der E-Mail möglich
  - Aber: Performance beachten

## MTA (II)

---

- Hat Zugriff auf mehr Merkmale als der Client
  - Einliefernde IP-Adresse
  - Protokolldaten
  - Verhalten des einliefernden Rechners
- Hat auf jeden Fall Verbindung ins Internet zur Überprüfung von DNSBLs, SPF, DKIM, ...

# Antispam-Maßnahmen

---

- DNSBLs (DNS Blacklists)
- SPF und SenderID
- DKIM (DomainKeys/IIIM)
- Prüfsummendatenbanken (DCC, Razor, Pyzor)
- Greylisting

# DNSBLs

---

- Listen mit IP-Adressen von Spammern
- Von den meisten Rechnern kommt nur Spam oder nur Ham
- DNS wird zur Speicherung benutzt
- Sehr einfach und billig
- Viele Betreiber solcher Listen, viele Policies
- Kann 20-40% des Spams erkennen

# SPF und SenderID

---

- Im DNS wird angegeben, von welchen Rechner E-Mail für eine bestimmte Domain kommen darf
- Eigentlich kein Antispam-Verfahren, soll nur Adressfälschungen verhindern
- Viel Presse, viel Hype, viele Probleme, viel Konfusion
- Funktioniert nicht wirklich

# DKIM

---

- Inhalt und Teile des Headers werden beim Versand signiert
- Den passenden Public Key gibt es im DNS
- Rechenaufwändig!
- Hat nicht so viele Probleme wie SPF/SenderID
- Standardisierung in Arbeit

# Prüfsummendatenbanken

---

- (Unschärfe) Prüfsumme über eine Mail
- Vergleich mit zentraler Datenbank
- DCC, Razor, Pyzor
- Wird in vielen kommerziellen Systemen verwendet

# Greylisting

---

- Ablehnen aller E-Mails beim ersten Zustellversuch
- Verzögerung der Zustellung
- False-Positives durch fehlerhafte Software
- Wird mehr und mehr verwendet...
- ...ist aber vielleicht schon bald wieder überholt

# MTA Software (I)

---

- Sendmail
- Exim
- Postfix
  
- Haben alle inzwischen umfangreiche Antispam-Features eingebaut und Schnittstellen für externe Hilfsprogramme

## MTA Software (II)

---

- Eingriffsmöglichkeiten in jeder Phase von SMTP: Connect, HELO/EHLO, MAIL FROM, RCPT TO, DATA
- Sendmail: Mail Filter API (“milter”)
- Exim: exiscan (seit 4.50 eingebaut)
- Postfix: Policy-Server
- Glue-Software: AmaViS, MailScanner, MIMEDefang, ...

# AMaViS

---



- “A Mail Virus Scanner”
- Aktuell: amavisd-new
- Kümmert sich um das Auspacken von MIME-Attachments usw.
- Kann Antiviren- und Antispam-Software einbinden
- Wird typischerweise vom MTA aufgerufen

# Procmail

---

- Filterprogramm, das typischerweise bei der Auslieferung einer E-Mail angewendet wird
- Verteilung von E-Mails in verschiedene Postfächer und dergl.
- Sehr flexibel
- Kryptisches Konfigfile
- Nicht nur für Spamfilterung

# SpamAssassin

---



- Sehr mächtig und flexibel durch Punktesystem (“scoring”)
- Weit verbreitet
- Auch Basis viele kommerzieller Produkte
- Relativ langsam
- Viele Antispam-Maßnahmen integriert...

# Funktionen von SpamAssassin

---

- Musterbasierte Erkennung
- Bayes-Filter (mit Autolearning)
- DNSBLs, URIDNSBLs
- DCC, Razor, Pyzor
- SPF
- Autowhitelisting
- ...

# dspam

---

- Statistischer Filter (Bayes u.a.)
- Als Filter (SMTP-Gateway) oder als Library
- Wird auch für kommerzielle Produkte verwendet
- Kommandozeilen- und Web-Interface
- Höhere Geschwindigkeit als SpamAssassin

# Bogofilter

---



- Bayes-Verfahren
- Typischerweise bei der Auslieferung der E-Mail eingesetzt
- Kann MIME-Attachments untersuchen

# CRM114

---



- Sehr flexibles Baukastensystem
- Basiert auf eigener Programmiersprache
- Filterung anhand von Mustern, Bayes, u.a.
- Nicht nur gegen Spam, auch zur Filterung von Logfiles etc.

# ClamAV Viren-Filter

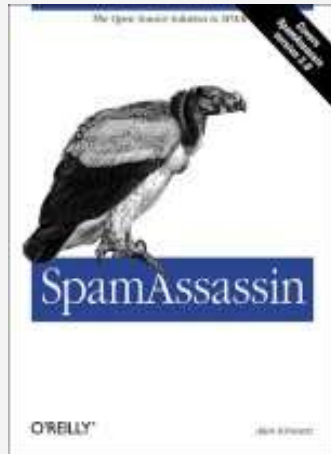
---



- Viren-Erkennung
- Einbettung in MTA, AMaViS, ...
- Leistung und Aktualität mit kommerziellen Systemen vergleichbar
- Erkennt 40.000 Viren, Würmer und Trojaner
- Automatischer Update ("freshclam")

# Bücher

---



Alan Schwartz:  
SpamAssassin.  
O'Reilly Media,  
2004



Peter Eisentraut,  
Alexander Wirt:  
Mit Open Source-  
Tools Spam &  
Viren bekämpfen.  
O'Reilly Verlag,  
2005



Alistair McDonald:  
SpamAssassin.  
Addison-Wesley,  
2005



Antispam-  
Strategien.  
Bundesanzeiger  
Verlag, 2005

# Weitere Informationen

---

- Weitere Bücher:  
<http://ref.AllThingsEmail.org/books/>
- Startpunkt Wikipedia:  
<http://en.wikipedia.org/wiki/Spamming>
- „everything you didn't want to have to know about spam“  
<http://spamlinks.net/>

# Noch Fragen?

---



Jochen Topf – [www.jtic.de](http://www.jtic.de) – [jochen.topf@jtic.de](mailto:jochen.topf@jtic.de)