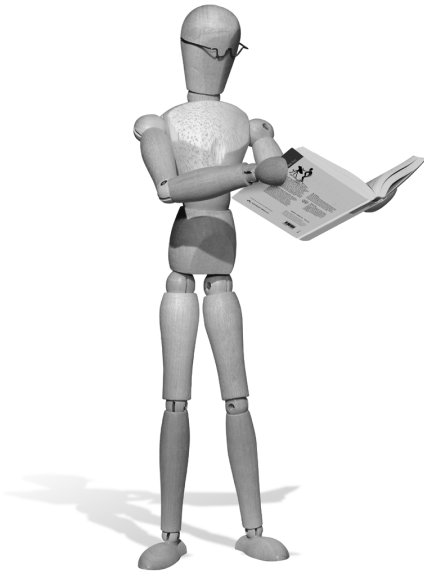


Alistair McDonald

SpamAssassin

Leitfaden zu Konfiguration, Integration
und Einsatz



ADDISON-WESLEY

An imprint of Pearson Education

München • Boston • San Francisco • Harlow, England
Don Mills, Ontario • Sydney • Mexico City
Madrid • Amsterdam

Inhaltsverzeichnis

Vorwort	17
Einleitung	19
Der Inhalt dieses Buchs	19
Was Sie für dieses Buch benötigen	21
Schreibweisen in diesem Buch	21
Die Beispiele dieses Buchs herunterladen	22
1 Einführung in das Thema Spam	23
1.1 Was ist Spam?	23
1.1.1 Definitionen	24
1.1.2 Spam im historischen Rückblick	24
1.1.3 Spammer	25
1.2 Die Kosten von Spam	25
1.2.1 Kosten für den Spammer	26
1.2.2 Kosten für den Empfänger	27
1.3 Spam vor dem Gesetz	27
1.4 Zusammenfassung	29
2 Spam- und Anti-Spam-Techniken	31
2.1 Spam-Techniken	31
2.1.1 Ausnutzung von offenen Relays	31
2.1.2 Sammeln von E-Mail-Adressen	32
2.1.3 Verschleierung des Inhalts	32
2.1.4 Statistisches Filter-Poisoning	32
2.1.5 Generierung einmaliger E-Mail-Texte	32
2.1.6 Trojanische Pferde	32
2.2 Anti-Spam-Techniken	33
2.2.1 Schlüsselwortfilter	33
2.2.2 Open Relay Blacklists (ORBLs)	33
2.2.3 Beschwerde beim ISP	33
2.2.4 Statistische Filterung	34
2.2.5 Analyse des E-Mail-Headers	34
2.2.6 Inhaltsüberprüfung	34

2.2.7	Whitelists	35
2.2.8	Inhaltsdatenbanken	35
2.2.9	Absendervalidierungssysteme	35
2.2.10	Sender Policy Framework (SPF)	36
2.2.11	Grey Listing	36
2.3	Spam-Filterdienste	36
2.3.1	Sammeln und Weiterleiten	37
2.3.2	Sammeln und Zurückschicken	37
2.3.3	Senden und Weiterleiten	37
2.3.4	Auswahl eines Anti-Spam-Dienstleisters	38
2.3.5	Dienste des ISPs	39
2.4	Anti-Spam-Programme	39
2.4.1	SpamAssassin	39
2.4.2	Zusammenfassung	41
3	Offene Relays	43
3.1	E-Mail-Zustellung	44
3.2	Testverfahren für offene Relays	44
3.2.1	Automatisierte Tests	45
3.2.2	Manuelle Tests	46
3.3	MTA-Konfiguration	47
3.3.1	Sendmail	47
3.3.2	Postfix	48
3.3.3	Exim	49
3.3.4	qmail	51
3.4	Zusammenfassung	51
4	E-Mail-Adressen schützen	53
4.1	Websites	53
4.1.1	Alternative Zeichendarstellung	54
4.1.2	JavaScript	54
4.2	Usenet	55
4.2.1	Trojanische Pferde	56
4.3	Mailinglisten und Archiv	56
4.4	Registrierung auf Websites	57
4.4.1	Die Verwendung von E-Mail-Adressen nachverfolgen	57
4.4.2	Unlautere Angestellte	58

4.5	Personal	58
4.6	Visitenkarten und Werbematerial	59
4.7	E-Mail-Validierung durch Spammer	59
4.7.1	Webbugs	60
4.8	Zusammenfassung	60
5	Spam erkennen	61
5.1	Inhaltsprüfung	61
5.2	Headerprüfung	62
5.3	DNS-basierte Blacklists	62
5.4	Statistische Tests	63
5.5	Nachrichtenerkennung	64
5.6	URL-Erkennung	64
5.7	Header untersuchen	65
5.7.1	Gefälschte Header	66
5.8	Spammer melden	67
5.9	Rechtmäßiger Versand von Massen-E-Mails	68
5.10	Zusammenfassung	70
6	SpamAssassin installieren	71
6.1	Erstellung aus dem Quellcode	72
6.1.1	Vorbereitungen	72
6.1.2	Ist ein C-Compiler vorhanden?	75
6.2	Verwendung von CPAN	76
6.3	Manuelle Installation	77
6.4	Build-Fehler reparieren	78
6.5	Distributionen	79
6.5.1	RPM	80
6.5.2	Debian	80
6.5.3	Gentoo	81
6.5.4	Andere Formate	81
6.6	Windows	81
6.7	Die Installation überprüfen	82
6.8	Aktualisieren	83

6.9	Deinstallieren	84
6.9.1	Deinstallation bei Verwendung des Quellcodes	84
6.9.2	Andere Pakete	85
6.9.3	Deinstallation unter Windows	85
6.10	Die Komponenten von SpamAssassin	85
6.10.1	Ausführbare Dateien	86
6.10.2	Perl-Module	86
6.10.3	Dokumentation	86
6.11	Zusammenfassung	86
7	Konfigurationsdateien	89
7.1	Konfigurationsdateien	89
7.1.1	Standardkonfiguration	89
7.1.2	Site-umfassende Konfiguration	89
7.1.3	Benutzerspezifische Konfiguration	90
7.2	Regeldateien	90
7.2.1	Regeln	90
7.2.2	Wertung	91
7.3	Zusammenfassung	93
8	SpamAssassin im Einsatz	95
8.1	SpamAssassin als Daemon	96
8.1.1	Benutzerkonten erstellen	97
8.2	SpamAssassin und Procmail	98
8.2.1	Ist Procmail vorhanden?	98
8.2.2	Procmail beschaffen und installieren	99
8.2.3	Procmail konfigurieren	99
8.2.4	MTA-Konfiguration	99
8.2.5	Benutzerkonten konfigurieren	102
8.2.6	Site-umfassender Einsatz von Procmail	104
8.3	SpamAssassin in den MTA integrieren	104
8.3.1	Sendmail	104
8.3.2	MIMEDefang	105
8.3.3	Postfix	107
8.3.4	Exim	107
8.3.5	qmail	108

8.4	Test und Fehlerbehebung	109
8.4.1	Überprüfen des MTA	110
8.4.2	Weiterführende Diagnose	111
8.5	Spam zurückweisen	112
8.6	Zusammenfassung	116
9	Bayes-Filterung	117
9.1	Wertung	117
9.2	Training	119
9.3	Ist der Filter aktiv?	120
9.4	Filtertraining	121
9.4.1	Benutzereingriff	121
9.4.2	Lokale Benutzer	122
9.4.3	Verlernen	123
9.4.4	Schwellenwerte für den automatischen Lernvorgang	124
9.4.5	Bayes-Datenbankdateien	124
9.4.6	Eine Bayes-Datenbank entfernen	125
9.4.7	Eine Bayes-Datenbank gemeinsam nutzen	125
9.5	Die Bayes-Filterung deaktivieren	127
9.6	Zusammenfassung	127
10	Benutzerdefinierte Anpassungen	129
10.1	Header	129
10.1.1	Header ändern	132
10.1.2	Header erstellen	132
10.1.3	Header entfernen	132
10.2	Berichte	133
10.2.1	Berichte aktivieren und deaktivieren	134
10.2.2	Berichte ändern	134
10.3	Den Betreff neu schreiben	136
10.4	Zusammenfassung	136
11	Netzwerktests	137
11.1	RBLs	139
11.2	SURBLs	140
11.2.1	SpamAssassin 2.63	141

11.3	Vipul's Razor	141
11.3.1	Razor installieren	142
11.3.2	Razor konfigurieren	142
11.3.3	SpamAssassin konfigurieren	145
11.3.4	Razor testen	146
11.4	Pyzor	147
11.4.1	Pyzor installieren	148
11.4.2	Pyzor konfigurieren	148
11.4.3	SpamAssassin konfigurieren	148
11.4.4	Pyzor testen	149
11.4.5	Pyzor-Header	150
11.5	DCC	150
11.5.1	DCC installieren	151
11.5.2	SpamAssassin konfigurieren	151
11.5.3	DCC testen	152
11.5.4	DCC-Header	153
11.6	Spam-Fallen	153
11.6.1	Adressen für Spam-Fallen	154
11.6.2	Köder legen	154
11.6.3	Das E-Mail-Konto konfigurieren	155
11.7	Zusammenfassung	156
12	Regeln	157
12.1	Regeln verfassen	158
12.1.1	Leistung von Regeln	162
12.1.2	Metaregeln	162
12.1.3	Positive Regeln schreiben	164
12.1.4	Rohtextregeln	166
12.1.5	Ein E-Mail-Corpus zum Testen der Regeln und der Wertung	167
12.2	Andere Regelsätze verwenden	170
12.3	Zusammenfassung	171
13	Filteroptimierung	173
13.1	Whitelists und Blacklists	173
13.1.1	Whitelists und Blacklists manuell anlegen	174
13.1.2	Domains in Whitelists aufnehmen	175
13.2	Die automatische Whitelist	176

13.3	Falsche Klassifizierungen berichtigen	177
13.3.1	Nachrichten untersuchen	178
13.3.2	Den Schwellenwert für Spam ändern	178
13.3.3	Testwertungen neu gewichten	180
13.3.4	Bayes'sches Verlernen und Neulernen	183
13.4	Zeichensätze und Sprachen	184
13.4.1	Sprachen ausschließen	184
13.4.2	Zeichensätze ausschließen	185
13.5	Zusammenfassung	187
14	Leistung	189
14.1	Engpässe	189
14.1.1	Speicher	189
14.1.2	Festplatten-E/A	191
14.1.3	Engpässe ermitteln	192
14.2	Methoden zur Leistungssteigerung	192
14.2.1	Den SpamAssassin-Daemon verwenden	195
14.2.2	SpamAssassin in den MTA integrieren	195
14.2.3	Nachrichten überspringen	195
14.2.4	Einige Testverfahren deaktivieren	197
14.2.5	Netzwerkbasierte Tests vorziehen	197
14.2.6	Zusätzliche Rechner verwenden	198
14.2.7	Schnellere Dateisperren	200
14.3	SQL verwenden	200
14.3.1	Voraussetzungen	201
14.3.2	MySQL	202
14.3.3	Spamd mit SQL	203
14.3.4	SQL für Benutzervoreinstellungen	203
14.3.5	SQL für Bayes-Datenbanken	208
14.3.6	Die Datenbank für die automatische Whitelist	210
14.4	Zusammenfassung	211
15	Wartung und Berichte	213
15.1	Spam nach Wahrscheinlichkeiten trennen	213
15.2	Fehler von SpamAssassin erkennen	214

15.3	Spam- und Ham-Berichte	216
15.3.1	Spam-Zähler	216
15.3.2	Die Verarbeitungszeit von SpamAssassin bestimmen	219
15.4	Zusammenfassung	222
16	Aufbau eines Anti-Spam-Gateways	223
16.1	Die PC-Plattform auswählen	224
16.2	Die Linux-Distribution auswählen	225
16.2.1	Linux installieren	226
16.3	Postfix konfigurieren	227
16.3.1	E-Mail an die Domain akzeptieren	228
16.3.2	Mail an den Benutzer root	228
16.3.3	Grundlegende Spam-Filterung mit Postfix	229
16.3.4	E-Mails an den ursprünglichen E-Mail-Server weiterleiten	229
16.3.5	Postfix neu laden	230
16.3.6	Postfix testen	230
16.4	Amavisd-new installieren	231
16.4.1	Installation von einem Paket	232
16.4.2	Vorbereitung	232
16.4.3	Installation aus dem Quellcode	233
16.4.4	Ein Benutzerkonto für Amavisd-new anlegen	233
16.5	Amavisd-new konfigurieren	234
16.6	Postfix für Amavisd-new konfigurieren	235
16.7	Externe Dienste konfigurieren	236
16.8	Die Firewall konfigurieren	236
16.9	Backups	236
16.10	Tests	236
16.11	Im Einsatz	237
16.12	Zusammenfassung	238
17	E-Mail-Clients	239
17.1	Allgemeine Konfigurationsregeln	239
17.2	Microsoft Outlook	240
17.3	Microsoft Outlook Express	245
17.4	Mozilla Thunderbird	247

- 17.5 Qualcomm Eudora 250
- 17.6 Zusammenfassung 251
- 18 Andere Anti-Spam-Programme 253**
 - 18.1 Spam-Richtlinien 253
 - 18.2 Spam-Filter bewerten 254
 - 18.3 Einen zweiten Filter konfigurieren 255
 - 18.3.1 Einzelner Rechner 255
 - 18.3.2 Getrennte Rechner 256
 - 18.4 Andere Techniken 259
 - 18.4.1 Greylists 259
 - 18.4.2 SPF 259
 - 18.4.3 Absendervalidierung 260
 - 18.5 Zusammenfassung 261
- 19 Glossar 263**
 - Stichwortverzeichnis 267**
 - Über den Autor 273**
 - Über die technischen Gutachter 275**