Mail-Server: SpamAssassin: jetzt mit Bilderkennung

Problem:

Seit geraumer Zeit vermehren sich Bilder in SpamMails. Meist sieht eine Mail so aus, daß ein (animiertes) Gif-Bild in normalen Text eingebettet wird.

Wer in der `local.cf` den Parameter `ok_languages` lediglich auf 'de' gesetzt hat, wird zumindest einen Teil der Spams filtern.

Lösung:

Besser fährt man allerdings mit einem kleinen Plugin für SpamAssassin: FuzzyOcr (als Erweiterung von OcrPlugin).

Vorbemerkung:

Seit Version 2.3 benötigt FuzzyOcr SpamAssasin ab der Version 3.1.4. Es wurde zwar versucht das Programm abwärtskompatibel zu halten, was aber überhaupt nicht funktionier.

Wer eine 2.0 Version hat, sollte wirklich mal updaten.

Wer eine 3.0 Version hat (z.B. mit Debian Sarge), der sollte entweder updaten (Debian-Stichwort: `volatile-sloppy`) oder die Version FuzzyOcr 2.2beta zurückgreifen.

Vorbereitung für FuzzyOcr:

Das Plugin braucht einiges an Software für den Betrieb. Dennoch fangen wir erst mit dem Plugin an: (aktuelle Version suchen!)

```
cd /usr/local/src/
wget http://users.own-hero.net/~decoder/fuzzyocr/fuzzyocr-latest.tar.gz
tar xzf fuzzyocr-latest.tar.gz
cd FuzzyOcr-*
```

Hier lesen wir uns erstmal die Datei `INSTALL` durch. Vorallem die Notes nach 1.6 für die verschiedenen Linux-Distributionen.

Danach machen wir es so wie es dort steht. Für Debian wäre dies z.B.:

```
apt-get install netpbm imagemagick libungif-bin
```

Die Perl-Module installieren wir über CPAN:

(c) 2025 huschi <huschi@huschi.net>
URL: https://www.huschi.net/index.php?action=artikel&cat=11&id=207&artlang=de

```
perl -MCPAN -e shell
install Digest::MD5
install String::Approx
```

Installation:

```
#Achtung: bei Debian gilt /etc/spamassassin/
cp -p FuzzyOcr.* /etc/mail/spamassassin/
cd /etc/mail/spamassassin
mv FuzzyOcr.words.sample FuzzyOcr.words
#evtl. müssen noch die Pfade in der FuzzyOcr.cf angepaßt werden.

#Prüfe die SA-Version (wenn latest-Version):
spamassassin -V
#Wenn kleiner als 3.1.4 muß folgende Zeile in der FuzzyOcr.cf gesetz werden:
focr_pre314 1

#Logfile zum Testen anlegen:
touch FuzzyOcr.log
chmod 666 FuzzyOcr.log

#SpamAssassin neu starten:
/etc/init.d/spamassassin restart
#Anlauffehler sind meist im maillog zu finden.
```

Testen:

Der Autor hat in der 2.3-Version netterweise bereits ein paar Beispiele mitgeliefert. Einfach mal durchprobieren:

```
spamassassin -t ‹ /usr/local/src/FuzzyOcr-2.3b/samples/jpeg.eml
```

Eindeutige ID: #1201
huschi
2006-10-23 13:32

(c) 2025 huschi <huschi@huschi.net>
URL: https://www.huschi.net/index.php?action=artikel&cat=11&id=207&artlang=de