SPAM

Projektarbeit 5. Semester

Martin Hausmann

E-Mail existiert seit ungefähr 20 Jahren, doch im Gegensatz zu anderen Technologien im Bereich des Internets hat sich E-Mail kaum verändert. Zum Senden wird immer noch das Simple Mail Transport Protocol (SMTP, RFC 2821) verwendet. Man kann hier eine beliebige Absenderadresse eintragen und hat im Gegensatz zu den Empfangsprotokollen (POP, IMAP) auch keine Authentifizierung durch Benutzername und Passwort. Und das war von den Entwicklern auch so gewollt, weil man 1982 die Bedrohung durch Spam-, Phising-, und andere unerwünschte E-Mails nicht kannte. Mittlerweile gibt es zwar Erweiterungen in diese Richtung (SMTP-Auth), diese sind aber nur optional und werden kaum verwendet. Diese Arbeit soll sich mit der SPAM Plage und möglichen Wegen um E-Mail noch zu retten beschäftigen.

1. Was kostet Spam?

1.1 Wie viel Geld kostet Spam?

Laut dem Magazin „Facts“ verursacht ein Spammail Kosten in Höhe von 2,7 Cent. Das klingt zwar nicht besonders viel, wenn man aber davon ausgeht, dass jährlich etwa 375.000.000.000 Spammails verschickt werden, kommt man auf Kosten in Höhe von rund 10 Milliarden Euro. Diese Kosten beinhalten jedoch nur die der Mailempfänger, nicht die der Provider von Mailservern und der Firmen deren Mitarbeiter wichtige Mails übersehen und löschen.

Laut einer Studie von Nucleus Research entstehen für Top-Konzerne in den USA zusätzliche Mehrkosten von 1.934 Dollar pro Mitarbeiter. Damit hat sich der Wert gegenüber dem Vorjahr mit 874 Dollar mehr als verdoppelt. Durchschnittlich erhalten die Angestellten der Studie zufolge 29 unerwünschte E-Mails pro Arbeitstag, gegenüber 13 im vergangenen Jahr.

1.2 Wie viele Ressourcen verbraucht Spam?

Spam kostet aber nicht nur Geld, sondern auch Ressourcen. Der Internetanbieter AOL geht beispielsweise davon aus, dass 30% aller Mails Spam sind. Laut der Homepage des Westfälischen Anzeigers sogar 47%. Das heißt ca. ein Drittel bis zur Hälfte der Serverkapazitäten unnötigerweise belastet werden. Die DejaNews-Datenbank, die seit 1995 Newsgroup-Postings sichert, und somit einen enormen Wissensschatz darstellt, besteht vermutlich aus zwei Dritteln aus Spam. Würde dieser nicht sein, würden mit einem Schlag riesige Kapazitäten wieder frei.

Auch die Nutzer leiden unter der Spamflut, denn wenn man pro Tag ungefähr zehn bis 20 E-Mails löschen muss kosten das neben Zeit auch Volumen, dass, falls man keinen „Flatrate“ Internetzugang hat, zu einer Beeinträchtigung des Monatsbudgets führt.

2. Welche Mittel gegen Spam gibt es?

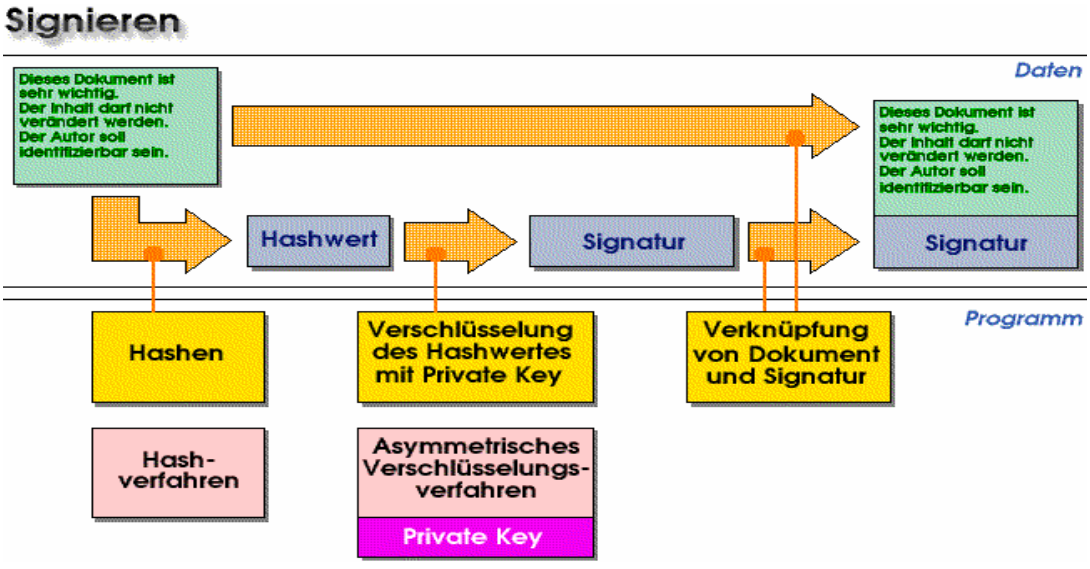
2.1 Digitale Signatur

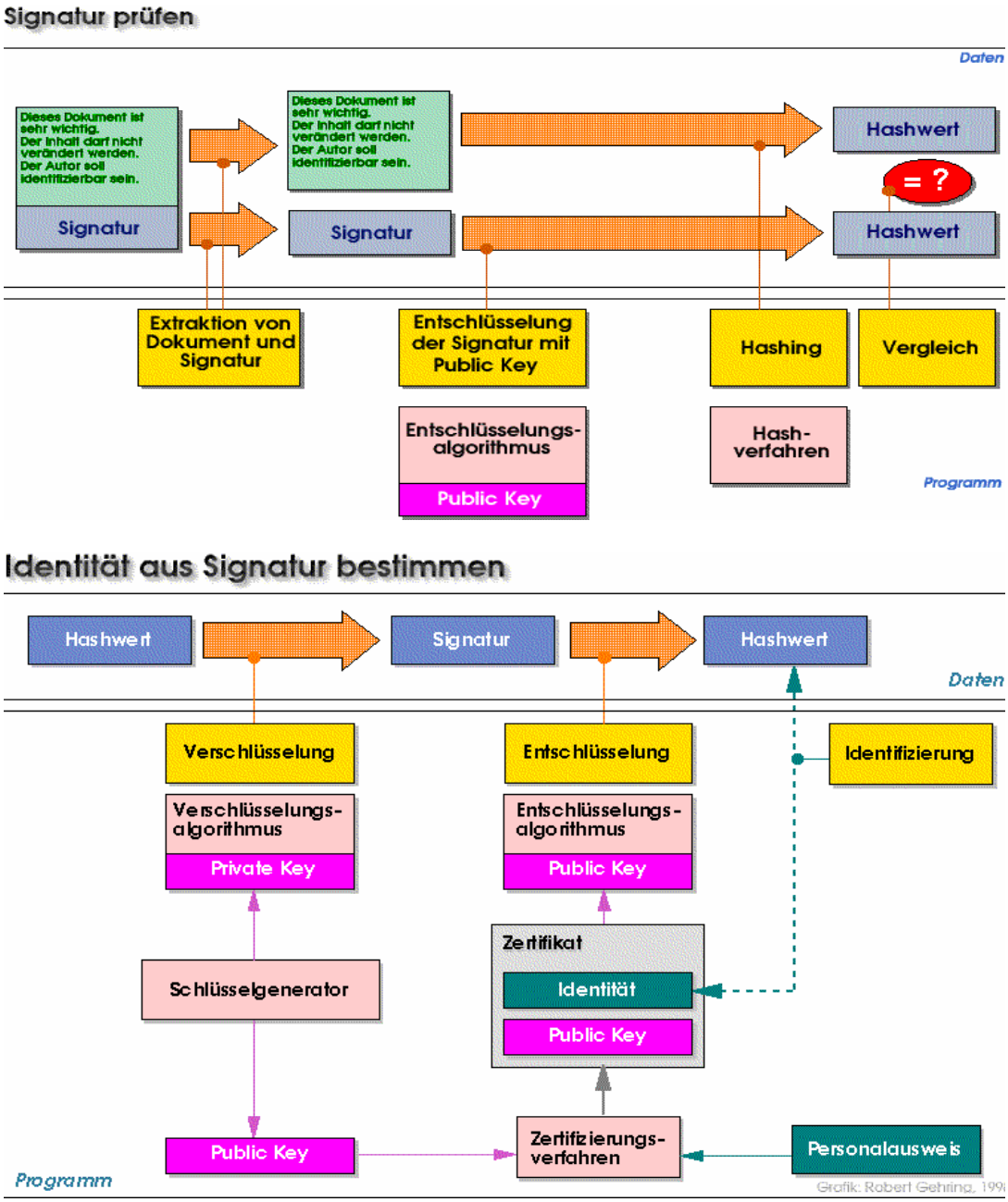
Die digitale Signatur wäre eine Lösung um Spam Einhalt zu gebieten. Man kann seine E-Mails zusätzlich zum – nicht fälschungssicheren – Absender auch mit einer Unterschrift versehen. Diese Möglichkeit gibt es schon seit einiger Zeit, wird aber kaum eingesetzt. Das hat durchaus seine Gründe.

Zum einem existieren zwei Standards die nicht kompatibel sind, und nicht jedes E-Mailprogramm unterstützt beide. Zum anderen kosten die benötigten Zertifikate teilweise Geld. Das führt dazu dass der Absender relativ viel Aufwand hat, um seine E-Mails zu signieren, während der Empfänger den Vorteil hat, dass er weiß von wem die erhaltenen E-Mails sind.

Nun zu den zwei Standards. Da gibt es einmal den S/MIME und auf der anderen Seite Open-PGP. Sie arbeiten zwar nach dem selben Prinzip, verwenden aber andere Datenformate, was dazu führt dass ein Open-PGP fähiges Programm keine S/MIME Signaturen lesen kann, bzw. nur mit einem PlugIn.

Das Signieren funktioniert nach folgendem Prinzip: Der User hat zwei Schlüssel, einen geheimen und einen öffentlichen. Versendet er nun eine Nachricht, erzeugt das E-Mailprogramm eine Prüfsumme des Nachrichteninhalts und hängt sie verschlüsselt an die E-Mail an. Der öffentliche Schlüssel wird mit übertragen und so kann der Empfänger mit diesem Schlüssel die Prüfsumme entschlüsseln und auf Korrektheit überprüfen. Stimmen sie überein, wurde die Nachricht mit dem geheimen Schlüssel, der zum mitgeschickten öffentlichen gehört, signiert und seither nicht verändert. Er hat so die Gewissheit, dass die Nachricht auch wirklich von dem Absender stammt der in der Absenderadresse steht. Nachfolgende Grafik der TU Berlin erklärt das Prinzip der Digitalen Signatur noch einmal bildlich.





2.2 SenderID

Durch den Vertrauensverlust der Benutzer sahen sich die Anbieter von Internetdiensten gezwungen ihre Rivalitäten beiseite zu schieben und zusammen eine Möglichkeit zu finden gegen Spam und Phishing vorzugehen. So haben AOL, Yahoo und Microsoft die „Anti Spam Technical Alliance“ (ASTA) gegründet und diese sucht nach Methoden um der Plage Herr zu werden.

Doch auch hier gibt es unterschiedliche Ansätze. AOL setzte schon früh auf das von Pobox entwickelte „Sender Policy Framework“, kurz SPF. Das SPF gilt als

aussichtsreichster Kandidat für die Standardisierung durch die „Internet Engineering Task Force“ (IETF), wenn auch in modifizierter Form.

Grundsätzlich geht es bei SPF, wie auch bei den anderen Methoden, um die Einbindung von Mail-Servern im DNS. Doch wie funktioniert SPF. Der versendende Mailserver holt sich beim DNS die IP des empfangenden Servers. Dieser wiederum versichert sich daraufhin, ob die Absenderangaben in der Nachricht („Envelope Form“) auch mit denen des liefernden Servers übereinstimmen. Die Entwickler von SPF stellen ihr Konzept als Open Source zur Verfügung.

Als Gegenkonzept von Microsoft, stellte Bill Gates die CallerID Methode vor. Der größte Unterschied zu SPF liegt darin, wie die DNS eingetragen wird. Bei SPF wie die DNS in Freitextzeichen eingetragen, wohingegen Microsoft auf XML setzt. Kritiker meinen allerdings, dass diese Methode einen XML fähigen Server voraussetzt und die XML Interpretation unnötigerweise die Serverperformance stört. Ein viel größerer Kritikpunkt liegt aber an der Tatsache, dass Microsoft ein Patent auf die XML-Syntax zum CallerID Verfahren angemeldet hat. So erklärt Microsoft auf der Homepage: Jeder darf einen CallerID Eintrag an seiner Maildomain vornehmen, Entwickler mögen aber bitte eine Lizenz beantragen, bevor sie eine Auswertungs-Engine für ihre Mailsoftware entwickeln.

Somit ist es nicht möglich CallerID in Open-Source-Software mit Gnu-Lizenz einzubauen und die Free Software Foundation warnt die IETF davor CallerID zu standardisieren, da Microsoft damit die ausschließt die sie ausschließen wollen.

Mittlerweile hat man sich aber auf eine Kombination zwischen SPF und CallerID geeinigt → SenderID. SenderID wird vermutlich innerhalb der nächsten Wochen das Rennen machen und zum RFC-Standard ausgerufen.

Ein Problem von SenderID ist aber das „Forwarding“. Da die Nachricht hier vom ersten Empfänger an den Zweiten weitergeleitet wird und somit nicht vom ursprünglichen Absender verschickt wird, würde sie am SPF-Filter hängen bleiben. Pobox hat hierfür das „Sender Rewriting Scheme“ (SRS) entwickelt, dass den E-Mail Header umschreibt und so den ursprünglichen Sender bei einem SPF Check ignoriert.

Einen wesentlichen Anteil am Erfolg von SenderID haben aber die großen Mailprovider, im deutschsprachigen Raum wären das etwa GMX, Web.de, T-Online oder AOL. Falls sie ihre Systeme wirklich so streng konfigurieren, dass alle Nachrichten die nicht mittels SenderID identifiziert werden können verworfen werden, dann müssen auch alle anderen Anbieter auf den Zug aufspringen.

Oder wenn sie das nicht tun, so befürchten Skeptiker, könnten zwei E-Mail Zonen entstehen. Die großen Provider schicken sich gegenseitig ihre validierten E-Mails und die kleinen Provider bleiben in einer Grauzone stecken. Dieses Szenario ist allerdings eher unwahrscheinlich, da es schon Tools und Serverpatches für SPF gibt.

2.3 IPv6

IPv4 hat ewig gehalten, aber mittlerweile gehen die IP-Adressen aus. Zum einen liegt das daran, dass IPv4 nur über 4 Mrd. möglicher Adressen (2^{32}) verfügt und heutzutage auch Handys, PDAs, Drucker, Spielekonsolen, Voice-over-IP-Telefone,… eine eigene IP benötigen. Ein weiteres Problem liegt darin, dass in den Anfängen des Internets ca. 70% aller IP-Adressen für die USA reserviert wurden, da man nicht davon ausging, dass sich das Internet so ausbreiten würde. So hat zum Beispiel die Universität Berkeley 16 Mio. IP-Adressen. Und eine Umstrukturierung der IP-Adressen würde zu viel kosten und das Problem der mangelnd vorhandenen IPs nur um einige Jahre verzögern.

Abhilfe soll nun IPv6 bieten. IPv6 verwendet 128 Bit lange Adressen, was die Anzahl der möglichen IPs auf 2^{128} erhöht. Die Adressen würden ausreichen um auf jeden Quadratmillimeter der Erde 600 Billiarden Hosts zu setzen.

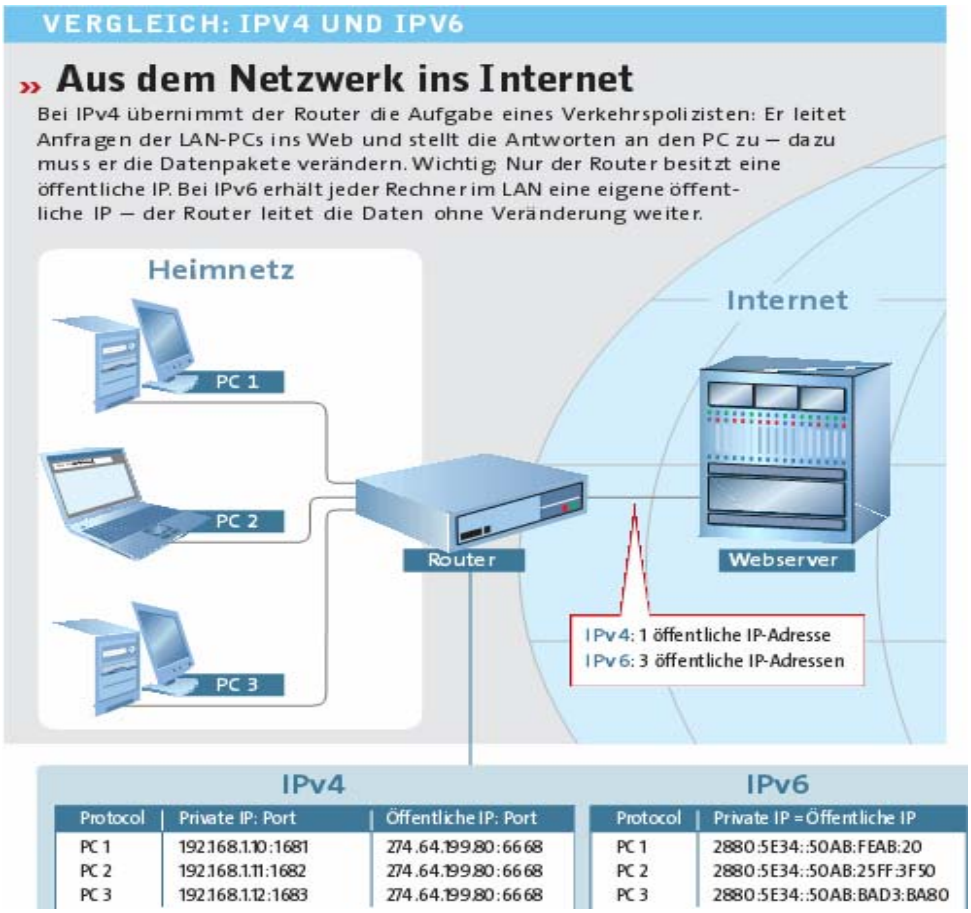
Zwar gibt es IPv6 schon seit 1995 aber in der Praxis findet man es äußerst selten. Durch die niedrigen Verkaufszahlen der letzten Jahre im Routerbereich nehmen aber immer mehr Anbieter die neuen Geräte ins Programm. Denn neben der Erhöhung der IP-Zahl, bietet IPv6 auch Erleichterung bei der Konfigurierung und fordert weniger Rechenleistung von den Routern.

Doch wie kann IPv6 gegen Spam helfen?

Durch den Mangel an IP-Adressen bei IPv4 hat nicht jeder PC in einem System eine eigene globale IP sondern eine interne und ist somit von Außen nicht mit seiner IP erreichbar. Durch die große Anzahl von IP-Adressen bei IPv6 könnte nun jedes Gerät eine eigene globale IP bekommen und wäre somit eindeutig identifizierbar.

Dadurch könnte man Spammern Einhalt gebieten, indem man bei E-Mails die IP des sendenden Rechners mitschickt. Die Implementierung einer solchen in die Programme würde keine besondere Herausforderung darstellen.

Das Problem daran könnte allerdings sein, dass die Spammer einen Weg finden, um diesen Mechanismus auszuschalten beziehungsweise zu umgehen.



2.4 Spammerlisten

Ein weiteres Mittel zur Spambekämpfung sind die Spammerlisten. Auf diese „schwarzen Listen“ kommen Provider von deren Systemen aus Spam verschickt worden ist. Ein gravierendes Problem ist aber, dass sobald ein Provider auf einer schwarzen Liste steht, unter Umständen sein Mailverkehr blockiert werden kann. Außerdem dauert es nach Behebung der Ursache oft mehrere Tage bis man wieder von der „Blacklist“ entfernt wird. Und die Spammer hält diese Maßnahme auch nicht dauerhaft auf, da sie einfach andere Server für ihre Zwecke missbrauchen.

Ein weiterer Punkt ist, dass nicht alle Provider auf dieselben Blacklists zugreifen, das heißt es gibt in Europa andere als zum Beispiel in den USA. Solange diese Differenzen nicht beglichen werden, wird auch diese Maßnahme nicht den gewünschten dauerhaften Erfolg bringen.

Eine etwas unorthodoxe Methode setzt nun Lycos Europe ein, in dem sie einen Bildschirmschoner anbieten der regelmäßig die Websites von Spammern aufruft mittels „http-GET-Requests“ (das bedeutet, dass keine Websites herunter geladen werden und somit auch kein auf der Website vorhandener gefährlicher Code, Dialer oder ähnliches mitkommen kann). Dabei werden Spammerlisten von „Spamcop“

zugrunde gelegt. Ziel ist es, bei den Spammern erheblichen Zusatztraffic und somit erhebliche Mehrkosten zu erzeugen. Laut Heise Online befindet sich das Ganze aber rechtlich gesehen in einer Grauzone.

2.5 Filtertechnologien

Spamfilter durchsuchen Mails nach bestimmten Kriterien und kennzeichnen, bzw. filtern sie.
Man unterscheidet hierbei nach der Art, wie die Filterprogramme die Mails untersuchen:

- Mails werden aufgrund bestimmter Kriterien - offensichtlich ungültige Absender, bekannte Spam Textpassagen, HTML-Inhalt, in die Zukunft datierte Absendedaten etc. – bepunktet und ab einer bestimmten Punktzahl als Spam klassifiziert.

- Es wird für jedes Mail ein Hash-Wert erzeugt und in Datenbanken überprüft, ob andere User das Mail erhalten haben und es als Spam klassifiziert haben.

- Eine andere Möglichkeit besteht darin, Mails mit dem Bayes-Filter zu kategorisieren. Dieser Filter lernt erst durch seinen Einsatz und setzt auf Worthäufigkeiten in bereits vom Benutzer erhaltenen und klassifizierten E-Mails. Diese Methode ist die am häufigsten eingesetzte und ist auch in vielen E-Mailprogrammen implementiert.

Die große Problematik hier ist aber, dass die Spammer sich diese Technologien zunutze machen und die Filter umpolen. Das heißt sie nutzen die Filter um Mails zu generieren die nicht den Kriterien entsprechen und somit auch nicht als Spam erkannt werden. So suchen zum Beispiel manche Filter vermehrt nach Frauennamen um so „Sex-Mails“ zu filtern. Um diese Hürde zu umgehen, verwenden die Spammer seltene Namen oder erfinden einfach Neue.

2.6 IBM

IBM geht es nicht um das „Neuerfinden“ vom E-Mail sondern um die Entwicklung von Werkzeugen die den Umgang mit den E-Mailmassen erleichtern sollen. Die Entwickler haben Anfang 2004 den Prototypen „REMAIL“ vorgestellt bei dessen Entwicklung und durch Forschung in den letzten zehn Jahren sich folgende Problembereiche herauskristallisiert haben:

- Anwender fühlen den Zwang, auf E-Mails schnell zu antworten. Ein ständiger Strom eingehender Nachrichten erhöht diesen Druck.

- Der Überblick über die E-Mails geht verloren und zugleich steigt die Furcht vor diesem Versagen. Große Mengen von Mails lassen die Wichtigen Nachrichten untergehen. Es wird unklar, welche Mails informellen Charakter haben, welche unwichtig sind und auf welche der Anwender reagieren muss.

- Die schiere Menge von Mails wirkt überwältigend. Um nicht ins Hintertreffen zu geraten, wird der Posteingang sehr häufig geprüft, was zu einer weiteren Produktivitätsverschlechterung führt.

Remail integriert Mail, Kalender und Chat. Sie macht es möglich seine Mails vollständig im Posteingang zu verwalten, bietet aber auch Collections genannte Ordner.

Sie lässt aber durch die Funktion „insight“ auch zu, seine Mails im Posteingang zu belassen und trotzdem in einem gewünschten Ordner zu lagern. Man kann noch zu erledigende Mails auch farblich unterlegen und zu einer Konversation gehörende Mails zusammen anzeigen lassen beziehungsweise hervorzuheben.

3. Conclusio

Meiner Meinung nach kann man Spam niemals ganz verhindern, da es sich bei dieser Problematik ähnlich verhält wie bei Kopierschutzmechanismen für CDs. Sobald ein neuer Schutz entwickelt und in Verwendung ist, dauert es oft nur kurze Zeit bis die Cracker einen Weg gefunden haben um den Mechanismus auszuhebeln.

Davon ausgehend, denke ich die beste Lösung wäre eine Mischung aus dem Einsatz der digitalen Signatur und SenderID. Man kann zwar mit der digitalen Signatur Spam nicht verhindern, aber man kann zumindest sicher gehen, dass die Nachricht auch wirklich von demjenigen kommt, der im Absender steht. Hierfür müsste aber ein einheitlicher Standard geschaffen werden, der von allen akzeptiert wird und vor allem gratis sein muss. Dies gilt hauptsächlich für die Nutzung im privaten Bereich, da die kommerzielle Nutzung immer eine andere Welt darstellt und sich, mit dem Verkauf von digitalen Signaturen an Firmen viel Geld verdienen lässt. Und diese Gewinne werden sich die Zertifkatsanbieter nicht entgehen lassen wollen.

Außerdem sollten Firmen, und da vor allem Banken und Kreditkartenfirmen, mit gutem Beispiel voran gehen und digitale Signaturen nutzen. Dadurch würden sie das Vertrauen ihrer Kunden in das Unternehmen stärken, und außerdem den „Phishern“ Einhalt gebieten. So wird zum Beispiel bei der Sparkasse Hannover seine Kreditkartenabrechnung als PDF File erhalten, ein Mail mit einem Link zur Downloadseite, auf der man seine Kreditkartennummer angeben muss. Das Mail ist

nicht signiert, aber die Seite hat ein Zertifikat, das aber leicht zu fälschen ist und somit sind „Phishern“ hier wieder Tür und Tor geöffnet.

Kombiniert man nun die vorhandenen digitalen Signaturen mit SenderID kann Spam doch weitgehend unterbinden, da durch die Überprüfung, ob der Sender überhaupt über diese IP senden darf, doch eine gute Methode darstellt sicher zu stellen, dass keine IPs missbraucht werden um Spam zu verschicken und man gleichzeitig sicher geht dass die Mails auch vom richtigen Absender kommen, verschließt man den Spammern und Phishern die Möglichkeiten zum Verschicken ihrer unerwünschten E-Mails.

4. Literaturverzeichnis

c’t Magazin Ausgabe 01/2000
c’t Magazin Ausgabe 16/2001
c’t Magazin Ausgabe 01/2004
c’t Magazin Ausgabe 10/2004
c’t Magazin Ausgabe 19/2004

CHIP Ausgabe 09/2004

http://ig.cs.tu-berlin.de/oldstatic/ap/rg/1998-06/abschnitt3.html 05.12.04

http://www.wa-online.de/computer/?storyid=1517 07.12.04

http://www.nucleusresearch.com/ 07.12.04

http://www.research.ibm.com/remail/ 30.11.04