Spambekämpning med SpamAssassin

Christopher Köppen

Magnus Lönnqvist

Kenneth Vinberg

2007

Kalmar, 2007-05-28
B-nivå, 5 poäng
Examensarbete datateknik
Handledare:
Marcus Wilhelmsson, Högskolan i Kalmar, Institutionen för Kommunikation och Design
Examinator:
Martin Kling, Högskolan i Kalmar, Institutionen för Kommunikation och Design

Institutionen för Kommunikation och Design
Högskolan i Kalmar

Sammanfattning

Under vårterminen 2007 genomförde vi ett examensarbete vid Högskolan i Kalmar. Arbetets syfte var att undersöka olika metoder för att minska mängden mottagen spam. För privatpersoner är det ett tidsödande göromål att skilja spam från legitim e-post då det i grund och botten är en subjektiv bedömning vad som egentligen är spam. En stor andledning till att skickande av spam görs möjligt är den låga kostnaden för att distribuera enorma mängder e-post. Det räcker i regel med att någon promille av mottagarna nappar på det erbjudande som avsändaren utannonserar, för att verksamheten ska löna sig.

Vi har valt undersöka en lösning baserad på programvaran SpamAssassin, vilken bygger på ett antal regler som används för bedömning av inkommande e-post. Utifrån dessa regler bedöms e-postmeddelandet enligt ett poängsystem som förtäljer huruvida e-postmeddelandet bör klassas som spam eller inte.

Inledningsvis simulerade vi en tilltänkt lösning på spamproblemet i virtualiseringsprogramvaran VMware. I ett senare skede satte vi även lösningen på prov i en praktisk tillämpning. I den pratiska tillämpningen uppnådde vi inte de proportioner spam kontra legitim e-post, samt förhållandet mellan text- och bildbaserad spam som enligt våra källor råder på Internet.

Resultatet av våra försök visar att SpamAssassin i sin grundkonfiguration inte uppnår hundraprocentig träffsäkerhet samt har stora svårigheter att identifiera bildbaserad spam.

Summary

During the spring term of 2007, we carried out our degree project at University of Kalmar. The purpose of the project was to examine different methods for minimizing the amount of received spam. For the individual it is a time consuming task to separate spam from legitimate email correspondence, as it is a subjective judgment to decide what is spam and what is not. A major reason for spam to be distributed is the low cost of sending large amounts of email at a very low cost. In general it takes just a few buyers per thousand recipients for the business to be profitable.

We have chosen to examine a solution based on the software SpamAssassin, which is based on a rule set used for assessing the likelihood of a given e-mail being spam.

Initially we simulated a proposed solution using the virtualization software VMware. At a later stage we chose to implement the solution in a practical environment. In the practical environment we did not reach the proportionate levels of received spam contra legitimate email that, according to our sources, currently exist on the internet.

The result of our experiment shows that SpamAssassin in its default configuration does not reach 100% accuracy and has major difficulties identifying image based spam.

Innehållsförteckning

1. Introduktion

Under vårterminen 2007 genomförde vi ett examensarbete vid Högskolan i Kalmar. Arbetet ämnade att undersöka möjligheten att filtrera spam med hjälp av mjukvaran Apache SpamAssassin. Denna rapport är en redogörelse för arbetets resultat samt lärdomar dragna av arbetet.

1.1. Problemformulering

För privatpersoner är det ett tidsödande göromål att skilja spam från legitim e-post då det i grund och botten är en subjektiv bedömning vad som egentligen är spam.

På arbetsplatsen blir spam ett problem då manuell genomgång av all e-post är väldigt tidsineffektivt, vilket direkt kan översättas i förlorad arbetstid och därmed pengar. På företagen blir spam även en kostnad på grund av mer och dyrare hårdvara för att hantera den mängd oönskad e-post som ständigt strömmar in.

Att ständigt översköljas med stora mängder information gör det svårt att se vilken information som är relevant, säker och tillförlitlig. Detta kan i sig bli ett stressmoment för den som använder e-post som arbetsredskap.

Stat och myndigheter i Sverige ställs inför ytterligare ett problem när det gäller spam, och det är offentlighetsprincipen. [1] Offentlighetsprincipen innebär att all post, inklusive e-post, som mottas av stat och kommun måste kunna offentliggöras. [2] Då inget av dagens e-postfilter uppnår en hundraprocentig träffsäkerhet kan man inte med automatik kasta någon e-post. [2] Därmed måste en anställd vid stat eller kommun först ha bedömt varifrån och från vem ett e-postmeddelande kommer, innan denne ges möjlighet att slänga det.

Trenden pekar stadigt uppåt vad gäller mängden spam kontra mängden legitim e-post. Enligt företaget Commtouch uppgick andelen spam till 85-90 % av den totala mängden e-post under slutet av första kvartalet 2007. [3] Flera andra företag och organisationer har kommit med liknande siffror. E-postfilterföretaget Postini publicerade i december 2006 en rapport där de uppgav att 94 % av all e-post som skickades i slutet av 2006 var spam. [4] Det bör i sammanhanget påpekas att Postini har ett visst egenintresse av att kunna uppvisa skrämmande statistik inom området. Postini pekar även på en markant ökning av bildbaserad spam, där hela meddelandet är uppbyggt av en eller flera bilder bifogade i e-postmeddelandet. De som skickar spam känner till att tekniken för att filtrera bildbaserad spam inte är lika väl utvecklad som för textbaserad spam, och har därmed högre sannolikhet att ta sig förbi automatiska e-postfilter. Bildspam blir även ett problem i och med att e-postmeddelandena blir större än dess textmotsvarighet, vilket innebär att mer bandbredd och lagringsutrymme går åt för att överföra och lagra e-postmeddelandet.

1.2. Syfte

Arbetets syfte är att undersöka olika metoder för att minska mängden mottagen spam. För att avgränsa arbetet inom detta omfattande område har vi valt att titta på lösningar baserade på öppen källkod och är kostnadsfria att implementera. Eftersom vi inte heller ämnar beröra specialanpassad hårdvara för e-postfiltrering eller externa betaltjänster så har vi valt att fokusera på programmet SpamAssassin och dess olika tekniker för spambekämpning.

2. Teori och teknik bakom e-post och spam

Här följer förklaringar av tekniska termer och programvara som läsaren behöver ha kunskap om för att kunna tillgodogöra sig rapportens innehåll.

Vi kommer att börja med att förklara vad e-post är och hur det fungerar, för att sedan komma in på hur spam definieras och vilken lagstiftning som gäller för spam. Slutligen behandlas programmet SpamAssassin och tekniker som används för identifiering av spam.

2.1. Hur fungerar e-post?

E-post skickas mellan e-postservrar enligt SMTP-protokollet som definieras i RFC 821 (se punkt 2.6.1 för förklaring av begreppet RFC). Vanligtvis skickar en klient meddelandet till sin lokala SMTP-server vilken kan tillhandahållas av avsändarens Internetleverantör eller företaget avsändaren jobbar vid. SMTP-servern i sin tur vidarebefordrar brevet till den mottagande e-postservern. Personen som brevet var adresserat till ansluter till sin e-postserver med IMAP4- eller POP3-protokollet för att läsa det mottagna meddelandet. [5, 6]

Hur ett e-postmeddelande utformas regleras av styrdokumentet RFC 2822. Detta bestämmer bland annat vilka fält som måste ingå i ett e-postmeddelande. Brevet är i huvudsak uppbyggt av ett brevhuvud och en meddelandekropp. Brevhuvudet består som minst av följande fält:

- Avsändare: E-postadress och eventuellt namn på avsändaren
- Mottagare: E-postadress(er) och eventuellt namn på mottagaren/mottagarna
- Ämne: Kort förklaring av meddelandets innehåll
- Datum: Lokalt datum och tidstämpel för när brevet skickades

Andra vanliga fält i brevhuvudet är:

- Cc: Carbon copy. Kopia till annan mottagare
- Bcc: Blind carbon copy. Hemlig kopia till annan mottagare
- Received: Spårningsinformation genererad av e-postserverar som hanterat brevet på vägen till mottagaren
- Content-Type: Information om hur meddelandet ska visas. Kallas vanligtvis ”MIME-typ”

Meddelandekroppen ska enligt RFC:n utgöras av US-ASCII tecken, med hjälp av olika MIME-typer kan meddelandekroppen formateras på många olika sätt, till exempel rena HTML-brev. Dessa MIME-typer behandlas i RFC 2046, 2048, 2049. [7]

2.2. Vad är spam?

Ordet SPAM (med versaler) är ett varumärke och det avser skinkkonserven Spiced Ham. Betydelsen av ordet spam härstammar från en sketch med Monthy Python. Sketchen går ut på att ett gäng vikingar sitter på en restaurang och sjunger ”spam, spam, spam, spam, spam”. De sjunger högre och högre, tills de andra gästerna får svårigheter att konversera. Det går här att dra en parallell till dagens spam genom att tänka sig att spam håller på att överrösta Internets alla e-postanvändare. [8]

Det finns ingen klar definition för vad spam egentligen är för något och begreppen kring detta är många. Spam, skräppost, obeställd e-post, obeställd e-post-reklam, bulkmail och ”unsolicited commercial bulk email” är några exempel på begrepp som brukar nämnas. [8]

I regel betyder ordet spam eller skräppost i e-postsammanhang ”obeställt massutskick”. Med obeställt menas att mottagaren inte uttryckligen givit tillåtelse för att meddelandet skickas till denne. Med massutskick menas att meddelandet är del av en större samling meddelanden med identiskt innehåll. [9]

Enligt organisationen Spamhaus (se punkt 2.9 för beskrivning av Spamhaus) definition räknas ett elektroniskt meddelande tekniskt sett som spam om det uppfyller följande villkor: [9]

- Mottagarens identitet och kontext är irrelevant då meddelandet kan gälla många andra mottagare.

- Mottagaren har ej uttryckligen givit tillåtelse för meddelandet att skickas till denne.

Nationalencyklopedins definition av spam lyder så här:

”spam (eng.), skräppost, massutskick på Internet av meddelanden som mottagaren vid fritt val skulle avstått från att erhålla, t.ex. reklam. Utskicken kan sändas till nyhetsgrupper eller till enskilda e-postmottagare med hjälp av listor över e-postadresser. […] Ibland liknas spam vid reklam via telefonnätet, men med skillnaden att sändaren av reklamen bär en mindre del av kostnaden för distributionen än den som tar emot reklamen” – Nationalencyklopedin, årsband 2003-10-22 [10]

Som mottagare kan du oftast inte tacka nej till fler utskick. Vid de tillfällen det finns möjlighet att tacka nej till mer e-post så har det ofta motsatt effekt eftersom avsändaren då får reda på att e-postadressen är giltig och används. [8] Spam har ofta ett kommersiellt budskap, där de lockar mottagaren med produkter eller kvasilegala tjänster. [11]

Spam är dock inte begränsat till den kommersiella kategorin, utan kan även placeras i kategorier så som: [12]

- Politiska meddelanden
- Välgörenhetsidkande
- Finansiellt/ekonomiskt bedrägeri
- Kedjebrev
- Meddelanden som sprider oönskad programvara
- Pornografi
- Hälsorelaterad marknadsföring

Avsändaren av spam är oftast anonym eller förfalskad. Kontakt med avsändaren sker då istället via en hemsida, vars adress står i meddelandet. [12]

Enligt organisationen Spamhaus blockeringsdatabas så skickas de flesta spammeddelandena från USA, Kina och Ryssland. [13] Enligt samma organisation kan 80 % av all skräppost riktad till mottagare i Nordamerika och Europa härledas till en grupp om cirka 200 personer som ägnar sig åt massutskick i stor skala. Dessa personers namn och alias finns listade på:
http://www.spamhaus.org/statistics/spammers.lasso

Det är värt att notera att samtliga av världens Internetleverantörer har bannlyst skickandet av spam, men effektiviteten i bekämpandet av spam varierar. [9]

2.2.1. Hur skickas spam?

Mottagarlistor kan exempelvis sammanställas genom att slumpa fram mottagaradresser till specifika domännamn. En enkel mjukvara kan till exempel slumpa fram ett förnamn, varpå detta kombineras med en lista över domännamn. [14]

En annan metod för adressinsamling är att läsa av publika webbplatser och nyhetsgrupper och automatiskt samla in publicerade e-postadresser. Ett program kan exempelvis samla in adresser från publika forum eller e-postlistor. [14]

Ytterligare en metod är att komma över en stulen databas med adresser från exempelvis en Internetleverantör eller webbtjänst. I spamkretsar förekommer även försäljning av adresslistor. [14]

Spam kan sedan skickas på ett antal olika sätt: [14]

1. Skickas direkt via egna eller hyrda servrar

2. Skickas via felaktigt konfigurerade servrar, vilka fungerar som öppna reläer.

3. Den vanligaste metoden är användandet av så kallade ”botnät”.

Ett botnät är en stor samling datorer vilka är infekterade med någon form av trojan. Detta botnät kan sedan styras centralt där en huvudman använder botnätet som plattform för att skicka ut spam utan datorägarnas vetskap eller tillåtelse.

En stor andledning till att skickandet av spam görs möjligt är den låga kostnaden för att skicka ut enorma mängder e-post. Det räcker i regel med att någon promille av alla mottagare nappar på det erbjudande som avsändaren utannonserar för att verksamheten ska löna sig. Vissa masspostföretag kan skicka ut 100 000 e-postmeddelanden för så lite som 2000kr. En databas med över en miljon e-postadresser kan köpas för under 1000kr. [2] Företaget Safe-Mail har ett räkneexempel där de menar att en spammare kan tjäna upp till tio miljoner amerikanska dollar per vecka, vid skickandet av 10 000 meddelanden för $1, förutsatt att denne har 100 000 kapade datorer till sitt förfogande. [15]

I Sverige och flera andra länder har det blivit allt vanligare att Internetleverantörer som tillhandahåller Internetåtkomst åt privatpersoner spärrar utgående trafik på TCP-port 25, vilken används av SMTP-protokollet för att skicka och ta emot e-post mellan e-postservrar. [16, 17, 18] I det fall en abonnents dator skulle bli infekterad av en trojan som försöker skicka spam, så kommer den trafiken på så vis blockeras redan hos den lokala Internetleverantören. Om kunden av någon anledning vill skicka e-post från sin egen e-postserver i hemmet kan denne använda Internetleverantörens SMTP-server som relä.

2.2.2. Lagstiftning

Från och med den 1 april 2004 är det i Sverige förbjudet att skicka e-postreklam som inte är beställd, och det finns idag liknande regler inom hela EU. [19] I huvudsak betyder detta att företag endast får skicka e-postreklam till personer och företag som i förväg tackat ja till att få reklam från företaget, eller om det finns ett etablerat kundförhållande mellan parterna. [20] Samtidigt infördes en definition i 3§ marknadsföringslagen för elektronisk post. Numer ingår förutom e-post även SMS- och MMS-meddelanden, fax och automatisk uppringning.

Sedan den 1 januari 2004 finns det i USA en lag vid namn CAN-SPAM (Controlling the Assault of Non-Solicited Pornography and Marketing Act). [21] Amerikanska Federal Trade Commission (FTC), vilket är motsvarigheten till det svenska Konsumentverket, är de som står bakom lagen och amerikanska justitiedepartementet har rätt att tillämpa lagen. [21] Lagen tillämpas för e-postmeddelanden vars syfte är att marknadsföra en kommersiell produkt eller tjänst, inräknat innehåll på hemsidor.

Sammanfattningsvis innebär lagen ett uttryckligt förbud för spammare att skicka missledande information, så som missledande ämnesrad, samt att de måste ge mottagare en möjlighet att kunna avböja fler meddelanden från spammaren via så kallad ”opt-out”. Vidare får spammare inte förfalska sina elektroniska avsändaradresser, och de måste dessutom i brevet bifoga en fysisk postadress på vilken de kan nås. [21]

Andra aktiviteter som enligt CAN-SPAM blir olaglig är bland annat automatiskt insamlande av e-postadresser från webbplatser och webbtjänster om de uttryckligen nekat insamling av e-postadresser i syfte att skicka e-post. [21] Automatiskt genererande av mottagaradresser, så kallade ordboksattacker, är inte heller tillåtet. Att reläa e-post genom annan e-postserver eller göra utskick från annans dator utan ägarens tillåtelse förbjuds också. Överträdelse av någon regel kan straffas med upp till $11000 i böter. [21]

Det finns dock en del kritik mot lagen. Enligt John Mozena, grundare av ”The Coalition Against Unsolicited Commercial Email” (CAUCE) så kommer vi knappast att bli av med spammarna. Lagen tvingar bara spammarna att vara hederliga i sitt spammande, genom att ange en riktig avsändaradress och ge möjlighet att exkluderas från framtida utskick. Vidare menar han att som lagen är utformad i dagsläget, där spammare får skicka spam tills dess att mottagaren uttryckligen begär att få slippa ta emot mer spam, är bakvänd. [22]

"It creates a situation where every legitimate marketer in the United States could send us one message. We could hit remove all day and not make a dent in the potentially tens of millions of possible marketers out there" – John Mozena 2003 [22]

Det har tagits upp som förslag att FTC ska upprätta en nationell opt-outlista, liknande det svenska NIX-registret för telemarketing. Även det förslaget är debatterat då det finns vissa svårigheter att se till att registret efterlevs samt den stora faran att den nationella opt-outlistan istället blir en nationell spamlista. [22]

2.3. Apache SpamAssassin

SpamAssassin är ett projekt som är skrivet med öppen källkod vilket innebär att programmets källkod är tillgänglig för allmänheten. [23] Tack vare detta kan programmerare enklare bidra till programmets utveckling, eller inspektera källkoden för att se hur programmet arbetar.

SpamAssassin påbörjades 1997. Redan från början byggde det på ett antal regler som jämfördes med inkommande e-post, och avgjorde utifrån ett poängsystem huruvida e-posten skulle levereras vidare till mottagaren eller inte. [23]

I sin nuvarande form använder sig SpamAssassin fortfarande av ett poängsystem för att avgöra om ett e-postmeddelande ska klassificeras som spam eller inte. [24] Ett antal kontroller utförs och resultatet av dessa ges enligt tidigare nämnda poängsystem. En matchande regel som tyder på att meddelandet kan vara spam ökar poängen, medan en matchande regel som innebär att e-posten verkar legitim sänker densamma.

När samtliga tester genomförts räknas poängen ihop till ett slutvärde vilket jämförs mot ett förbestämt tröskelvärde. Om poängen överskrider detta tröskelvärde markeras e-posten som spam genom att ett eller flera fält läggs till i brevhuvudet.

SpamAssassin analyserar både brevhuvud och meddelandekropp för att avgöra huruvida det mottagna e-postmeddelandet ska markeras som spam.

Vid analys av brevhuvudet söker SpamAssassin efter bland annat: [25]

- Ämnesraden innehåller medvetna felskrivningar av ordet ”Valium” eller ”Cialis” i syfte att förvilla spamfilter
- Saknar avsändaradress i ”Reply-to” -fältet
- Avsändarens IP-adress är listad i en RBL (Realtime Block List, se punkt 2.9)

I meddelandekroppen letar SpamAssassin bland annat efter vissa specifika ord och formuleringar, så som: ”Money back guarantee”, ”FREE ACCESS”, eller ”No Medical Exams”. SpamAssassin letar även efter adresser till kända spamsidor och hur pass stor mängd felformaterad HTML brevet innehåller. [25]

Reglerna som används kan modifieras efter egna behov. [26] För att ge ett exempel kan vi anta att man driver en möbelaffär. För att undvika att önskad e-post sållas bort av SpamAssassin kan egna regler läggas till där ord som ”stol”, ”bord” eller ”säng” ger minuspoäng i bedömningen.

För att ytterligare förbättra sin träffsäkerhet använder sig SpamAssassin av Bayesisk analys. Analysen är grundad på ett matematiskt teorem format av 1700-talsmatematikern Thomas Bayes. Detta teorem beskriver sannolikheten att en händelse kommer att inträffa baserat på hur många gånger händelsen inte inträffat vid tidigare försök. [27] Genom SpamAssassins implementation skapas ett register över ord och ordföljder, samt hur ofta dessa förekommer i önskad respektive oönskad e-post. Genom jämförelser mot detta register bedöms alltså sannolikheten att ett e-postmeddelande är oönskat baserat på formuleringar och ord som använts.

En egenhet hos den Bayesiska analysen är att ”inlärning” måste ske. Denna inlärning avser uppbyggnad av registret över ord och ordföljder som skall tas i beaktning i den Bayesiska analysen. Genom att användaren specificerar ett antal e-postmeddelanden som är legitima respektive spam kan denna process göras träffsäker. Om misstag görs under inlärningen kan den Bayesiska analysen resultera i felaktiga bedömningar och därmed ha en negativ effekt på filtrets träffsäkerhet. [27]

2.4. FuzzyOcr

FuzzyOcr är en tillägsprogramvara till SpamAssassin. FuzzyOcr fungerar som en bildavläsare som läser ut ord ur bilder, som sedan kan poängsättas enligt SpamAssassins regler. Programmet innehåller även funktioner för att räkna fram en kontrollsumma som innebär ett unikt id för en bild och jämföra denna kontrollsumma mot databaser med bilder kända för att förekomma i spam. [28]

Att enbart analysera ett e-postmeddelandes kontrollsumma är inte alltid tillräckligt, då spammarna ofta på automatiserat vis förändrar bilden med en eller flera bildpunkter, och på så sätt får en ny kontrollsumma. [4, 29]

2.5. Postfix

Postfix är en SMTP-server skriven av Wietse Venema. Målet med utvecklingen var att konstruera en SMTP-server som är snabb, säker och enkel att administrera. [30] Postfix är ett populärt alternativ till det vanligare, men mer svårkonfigurerade SMTP-servern Sendmail.

Postfix kan köras under operativsystemen AIX, BSD, HP-UX, IRIX, LINUX, MacOS X, Solaris, Tru64 UNIX, och andra UNIX-system. [31]

Det finns ett antal inställningar som kan göras i Postfix konfiguration för att begränsa mottagen spam redan på SMTP-nivå. Exempel på detta är:

- Kräv HELO eller EHLO-meddelande innan post mottas. [32]
- Fördröj svaret på HELO-meddelandet. Spammare har oftast inte tid att vänta särskilt länge, utan väljer istället att försöka leverera till annan e-postserver. [32]
- Kontrollera att avsändarens domännamn är ett giltigt domännamn. Detta förhindrar mottagande av e-post från falska eller påhittade domäner. [32, 33]
- Postfix kan även konfigureras för att göra uppslag mot Spamhaus blocklista redan vid första kontakt med sändande e-postserver. [32, 33]

2.6. RFC 821

2.6.1. Vad är en RFC?

En ”Request For Comment” (RFC) är en standard publicerad av Internet Engineering Task Force (IETF). IETF är en självorganiserad grupp människor som arbetar med Internets funktionalitet och utveckling. Målsättningen med arbetet som utförs av IETF är bland annat att identifiera och föreslå lösningar på problem i de tekniker som utgör Internets funktionalitet. [34]

Eftersom IETF inte är någon form av företag eller organisation utan en löst sammansatt grupp baserad på frivilligt deltagande kan vem som helst försöka driva igenom en RFC. Varje förslag granskas och revideras dock noggrant. [34]

2.6.2. Detaljer kring RFC 821

RFC 821 är namnet för standarden som beskriver Simple Mail Transport Protocol (SMTP). Denna standard fastställdes 1982 och författaren var Jonathan B. Postel, då verksam vid University of Southern California. I enlighet med IETFs höga krav beskrivs protokollet ingående och vetenskapligt. Bland annat beskrivs hur de operativa kommandona skall användas och hur ett överföringsförlopp ska gå till. [35]

År 2001 fastställdes RFC 2821, författad av J. Klensin vid AT&T Laboratories, som en utökning av RFC 821. Den nya RFCn innehåller tillägg, uppdateringar och förtydliganden till den ursprungliga, nu föråldrade standarden. [36]

2.7. Grålistning

Grålistning är en metod för att minska mängden mottagen spam som e-postservern behöver behandla. Metoden innebär minimal prestandapåverkan på e-postservern, då det är mycket små datamängder som behöver behandlas för varje e-postmeddelande. [37]

Det fungerar på så sätt att en ”triplett”, bestående av avsändarens IP-adress, avsändarens e-postadress och mottagarens e-postadress, registreras av den mottagande e-postservern. [37] Om det är första gången som den mottagande e-postservern stöter på en specifik triplett så kommer e-postservern svara den sändande e-postservern med ett tillfälligt felmeddelande. [37] Tripletten läggs sedan till den så kallade grålistan, där den registreras som tillfälligt blockerad i ett antal minuter. [37] SMTP-protokollets RFC 821 dikterar att den sändande e-postservern ska göra nya försök att sända meddelandet med jämna intervaller. [35] Detta är dock något som de flesta spammare väljer att inte göra, utan låter helt enkelt bli att göra ytterligare försök. Anledningen är att spammarna måste skicka så många spam som möjligt på kortast möjliga tid innan den avsändande IP-adressen registreras i blockeringslistor. [37]

Om källan däremot är legitim kommer den att försöka sända meddelandet på nytt. Tripletten är vid det här laget borttagen från bevakningslistan och kommer att läggas till en så kallad ”white-list”, där tripletten kommer att vara registrerad i cirka 30 dagar. Tidsgränsen nollställs varje gång tripletten återkommer. Detta för att exempelvis regelbundna nyhetsbrev inte ska fastna i grålistan. [37]

Enligt tidningen Ny Teknik utgiven 2005-03-02 så har Lunds universitet lyckats minska mängden mottagen spam med 90 % tack vare grålistning. [38] Det är ofta liknande siffror som presenteras, till exempel av skaparen av grålistning Evan Harris. På sin hemsida (http://projects.puremagic.com/greylisting/whitepaper.html) uppger han att man vid initial testning lyckats filtrera bort 95 % av all mottagen spam. [37]

Många är dock överens om att grålistning inte är en lösning som kommer att vara för evigt, utan kommer att tvinga spammare att börja följa RFC:erna och sända om e-postmeddelandena. Dock är det inte förgäves även om spammare skulle sända om e-postmeddelandena, eftersom en fördröjning ger större chans att det hinner detekteras som spam av anti-spamorganisationer och därmed registreras i blocklistor och hashdatabaser. [39]

2.8. Hashdatabaser

Spambekämpning kan med fördel kompletteras med användandet av distribuerade hashdatabaser. Dessa databaser innehåller kontrollsummor för kända spammeddelanden. Den mottagande e-postservern beräknar fram en kontrollsumma av det mottagna e-postmeddelandet, och skickar den till databasen för jämförelse. Databasens svar innehåller ett bekräftande eller nekande till om det finns registrerat samt hur många gånger kontrollsumman registrerats. Utifrån detta svar kan SpamAssassin välja att lägga till poäng i sin bedömning. Ett exempel på en dylik databas är Distributed Checksum Clearinghouse (DCC) (http://www.rhyolite.com/anti-spam/dcc/). Efter behandling av ett e-postmeddelande rapporterar DCC-klienten in kontrollsumman för meddelandet oavsett om det är spam eller inte. Därför är det en bra idé att upprätta en så kallad white-list som hindrar DCC från att rapportera in legitima nyhetsbrev. [40]

Razor2 är en annan hashdatabas som kan användas av SpamAssassin. Den fungerar på ett liknande sätt men med skillnad att rapportering till Razor2-databasen sker genom registrerade rapportörer, som över tid bygger upp en pålitlig relation beroende på hur pass tillförlitlig deras rapportering tidigare varit. [41]

Pyzor är ytterligare ett exempel på en distribuerad hashdatabas som används för uppslag vid poängsättning av inkommande e-postmeddelanden. [42]

SpamAssassin har stöd för uppslag mot hashdatabaserna DCC, Razor2 och Pyzor. [43, 44, 45] SpamAssassin lägger då till ytterligare poäng till bedömningen, beroende på svaren från hashdatabaserna.

2.9. Realtime Block Lists

Spamhaus är en internationell organisation vars mål är att spåra spammare på Internet, tillhandahålla realtidsuppdaterade anti-spam skydd, arbeta tillsammans med rättsväsenden för att stävja spam samt att organisera lobbyverksamhet för att etablera en fungerande anti-spamlagstiftning. [46]

Spamhaus SBL är en databas med IP-adresser tillhörande kända e-postservrar och webbservrar som förknippas med spamverksamhet. SpamAssassin kan använda denna databas för att avgöra om den sändande partens IP-adress är associerad med spamverksamhet, och kan därmed välja att lägga till ytterligare poäng om databasen rapporterar avsändaren som trolig spammare.

Spam innehåller i de flesta fall en adress till sändarens webbplats, där den utlovade produkten eller tjänsten marknadsförs. SpamAssassin kan extrahera dessa länkar från e-postmeddelandet och göra en förfrågan till Spamhaus SBL för att få reda på ifall länken leder till en webbplats associerad med spamverskamhet. [47]

IP-adresserna hamnar i Spamhaus blocklista via så kallade ”spam traps”. En spam trap är en mottagaradress som inte tillhör en riktig användare. Adressen är aldrig publicerad någonstans där en människa kan hitta den. E-post till denna adress kan omöjligen vara ombedd då ingen människa finns bakom den adressen, och därmed klassas den automatiskt som obeställd e-post. [48, 49] IP-adresserna finns kvar i databasen enligt en tidsgräns som definieras vid registrerandet. Det kan vara allt mellan två dagar till ett år. Vid registrering av större IP-segment så skickar Spamhaus ut en varning till segmentets ägare som exempelvis kan vara en Internetleverantör. [48, 49]

3. Metod

Inledningsvis simulerade vi en tilltänkt lösning på spamproblemet. Vi valde att simulera två e-postservrar i en VMware-miljö, där den ena är en mottagande e-postserver och den andra är en spamskickande e-postserver. Anledningen till att vi valde att simulera lösningen i en kontrollerad miljö var att vi då kunde styra mängden och urvalet spam som skulle bedömas.

Båda e-postservrarna installerades med GNU/Linuxoperativsystemet Ubuntu 6.10 Server, och valet av SMTP-server föll på Postfix. Vi valde att använda oss av GNU/Linux-operativet Ubuntu med anledning av att det finns ett stort urval av dokumentation för installation och konfiguration av dess tjänster. Valet av e-postserver föll på Postfix av den anledningen att den är förhållandevis enkel att konfigurera tillsammans med SpamAsssassin och att det finns gott om konfigurationsdokumentation.

Den mottagande e-postservern preparerades med Apache SpamAssassin och konfigurerades att använda DCC, Razor2 och Pyzor samt att använda det förinställda tröskelvärdet 5.0. På samma server installerades även IMAP-programvaran Courier-IMAP för att kunna läsa mottagen e-post. Vi valde att använda oss av tilläggsprogramvarorna Razor2, Pyzor och DCC som SpamAssassin har inbyggt stöd för, därför att det ger en mer träffsäker bedömning huruvida ett e-postmeddelande ska klassas som spam eller inte. Det ursprungliga standardtröskelvärdet 5.0 behölls för att kunna ge en bild av huruvida SpamAssassin i sin grundkonfiguration är för generös i sin bedömning eller inte.

Vi valde att inte använda oss av tilläggsprogramvaran FuzzyOcr med anledning av att SpamAssassin i sitt standardutförande inte har ett insticksprogram för användning av FuzzyOcr.

Postfix och SpamAssassin knöts samman med hjälp av ett mindre bashskript, för överlämning av inkommande e-post från Postfix till SpamAssassin. Skriptet ser ut som följande:

```
#!bin/bash
/usr/bin/spamc |/usr/sbin/sendmail –i ”$@”
exit $?
```

Skriptkod 3.0.1: Bashskript för överlämning av inkommande e-post till SpamAssassin

3.1. Genomförande i simulerad miljö

Med hjälp av en klientdator laddade vi ned en större mängd spam från en av projektdeltagarnas privata e-postkonto på gratistjänsten Google Gmail. Ett urval av dessa spam skickades sedan via den sändande e-postservern till den mottagande e-postserven som hade till uppgift att poängsätta varje mottaget e-postmeddelande.

Sedan kontrollerades de mottagna e-postmeddelandena för att undersöka huruvida de av SpamAssassin identifierats som spam eller inte. Detta har gjorts genom att undersöka e-postmeddelandenas ämnesrad efter omskrivningen som SpamAssassin gör vid identifiering av spam. Vi har även analyserat de mottagna e-postmeddelandenas brevhuvud, där vi kunnat utläsa antalet poäng som SpamAssassin tilldelat det aktuella e-postmeddelandet.

Enligt skaparna av programmet fungerar inte SpamAssassin som avsett i en simulerad miljö. Om man skickar e-post till sig själv, så är e-posten inte är oombedd samt att den inte är en del av ett massutskick. Därmed kan det enligt SpamAssassins skapare inte klassas som spam. [50] Vi valde ändå att genomföra ett simuleringsförsök för att bekanta oss med SpamAssassin och dess konfiguration.

3.2. Genomförande i praktiken

För att testa SpamAssassin under verkliga förhållanden registrerade vi domännamnet ”julpost.com”. Vi konfigurerade sedan en e-postserver med samma konfigurationsparametrar och mjukvaror som i den virtuella miljön. Denna server anslöts sedan till Internet för att mottaga e-post för julpost-domänen.

Projektmedlemmarna hade varsin e-postadress att disponera för mottagande av nyhetsbrev och annan legitim prenumerationsbaserad e-post.

E-postadressen olle@julpost.com spreds på diverse hemsidor, forum, nyhetsgrupper och andra publika webbplatser. Denna e-postlåda skapades för att mottaga all e-post som inte var adresserad till någon av projektmedlemmarna, en så kallad ”catch-all”-adress. Med hjälp av denna kunde vi sedan se vilka e-postadresser som genererats fram av spammare eller snappats upp på de ställen vi publicerat e-postadressen olle@julpost.com.

Insamlandet av e-post pågick under en 30-dagarsperiod, och efter denna period sammanställdes resultaten av försöket.

4. Resultat

4.1. Resultat av test i den virtuella miljön

Inledningsvis hade vi stora problem att få SpamAssassin att identifiera inkommande e-post som spam. Problemet var att den mottagande e-postservern till en början inte hade tillgång till Internet och därmed inte heller blocklistor och hashdatabaser. I och med detta fick SpamAssassin uteslutande förlita sig på enkel innehållsanalys av e-posten där den letade efter vanligt förekommande ord och formuleringar, så som ”No prescription needed”. Detta resulterade i att så gott som inga e-postmeddelanden fick tillräckligt höga poäng för att markeras som spam. Vår erfarenhet visar på att enbart innehållsanalys utan blocklistor eller hashdatabaser inte är tillräcklig för att göra en korrekt bedömning.

När vi sedan försåg den mottagande e-postservern med Internetåtkomst fick vi mycket bättre resultat. En övervägande majoritet av alla textbaserade spam fick nu tillräckligt höga poäng för att markeras som spam. Detta berodde på att den mottagande e-postservern nu kunde göra uppslag mot blocklistor och hashdatabaser.

För att kunna redogöra ett resultat av spamfiltreringen i den virtuella miljön skickades 60 stycken olika spam till den filtrerande e-postservern. E-posten som skickades var representativa för de proportioner bildbaserad och textbaserad spam som mottagits till våra privata e-postkonton. Observera att vi enbart skickat manuellt inspekterade och därmed verifierade spam till den mottagande e-postservern. I och med det fanns det inga legitima e-postmeddelanden som felaktigt kunnat markeras som spam.

Nedan redovisas resultatet i tabellform.

	Markerade	**Ej markerade**	**Totalt**
Bildspam	0	18	18
Textspam	20	5	25
Blandad bild/textspam	2	15	17
			Totalt 60

Tabell 4.2.1: Resultat av test i virtuell miljö

Tabellen visar att SpamAssassin i sin grundkonfiguration har stora svårigheter att korrekt markera bildbaserad spam. Spam med blandat text- och bildinnehåll var även problematiskt att identifiera korrekt. Textbaserad spam markerades emellertid med 80 % träffsäkerhet.

4.2. Resultat av test i praktiken

Under försöksperioden mottog vi totalt 448 e-postmeddelanden. Genom manuell inspektion verifierades 59 stycken som spam. Av dessa 59 meddelanden blev sju stycken ej markerade som spam. Sju legitima e-postmeddelanden markerades felaktigt som spam. Under försöksperioden mottog vi totalt två stycken e-postmeddelanden med fingerad mottagaradress.

Sammanställning av resultatet visar att cirka 13 % av e-posten som mottogs under testperioden var spam. Andelen e-post som felaktigt markerats som spam uppgick till ungefär 1,5 %. En lika stor andel e-post markerades felaktigt som legitim e-post, se tabell 4.2.2.

Spam totalt	Legitim e-post	Felaktigt markerad som spam	Felaktigt markerad som legitim e-post	Fingerad mottagaradress	Totalt
59	375	7	7	2	448

Tabell 4.2.2: Resultat av praktiska test

Av den spam som mottogs var nästan 19 % HTML-brev med invävda bilder. Resterande 81 % utgjordes av textbaserad spam. Se tabell 4.2.3.

HTML/bildspam	Textspam	Totalt
11	48	59

Tabell 4.2.3: Andelar text- respektive bildbaserad spam

5. Diskussion

I vårt praktiska försök uppnådde vi inte de proportioner spam kontra legitim e-post, samt förhållandet mellan text- och bildbaserad spam som enligt våra källor råder på Internet. Båda dessa missförhållanden i proportionalitet beror i huvudsak på det faktum att domännamnet som användes i försöken inte hunnit sprida sig till tillräckligt många spamkällor för att ge en bra helhetsbild.

Vi tror att vi kunnat uppnå ett bättre resultat om vi kunnat lära upp det Bayesiska filtret med hjälp av inkommen spam och legitim e-post, men det var tyvärr inte möjligt då detta kräver en större mängd e-post. Vidare tror vi att ett bättre resultat kunnat uppnås avseende klassning av bildspam om vi använt oss av ett insticksprogram för bildläsning, exempelvis FuzzyOcr.

De begränsningar vi valt att göra beror i största del på vårt intresse för kostnadseffektiva spambekämpningslösningar samt de möjligheter som ges av öppen källkod.

Vi hade önskat implementera denna lösning på en existerande domän för att få en större mängd data att behandla. I och med det hade vi kunnat dra en mer exakt och verklighetsförankrad slutsats om lösningens effektivitet. Det hade dessutom varit intressant att se vilken belastning e-postserverns hårdvara utsätts för vid filtrering av större volymer e-post.

5.1. Framtiden för spambekämpning

För att få ett övertag i den ständigt pågående kampen mot spammare finns ett antal förslag som kan bli aktuella i framtiden:

- **Strängare lagstiftning för skickande av spam**
  Om en teknisk lösning på spamproblemet inte tas fram så är det sannolikt att en internationell lagstiftning blir en nödvändighet för att stävja problemet. Som i alla fall när internationell lagstiftning ska införas så är det en tidskrävande process, samt att organisation för kontroll av lagens efterlevnad måste etableras.

- **Avgift för att skicka e-post**
  Inte helt enkelt att införa, och dessutom osäkert om spammare egentligen kommer att betala, då de oftast skickar e-posten med falsk avsändaradress. Därtill finns frågan om hur pengarna ska debiteras och vem som ska ta emot dem. [51]

- **Omskrivning av protokollen**
  Den mest radikala lösningen på problemet är en total omskrivning av protokollen som styr hur e-post skickas, men enligt Laura Atkins vid SpamCon Foundation kan ett sådant arbete ta flera år. [51] Ett exempel på vad som kan införas är krav på identifiering av avsändaren. Då skulle man på ett enkelt sätt kunna välja att blockera e-post från icke-identifierade avsändare. Denna idé strider dock mot det långsiktiga målet med ett öppet e-postsystem utan central hantering, vilket tillåter anonym yttrandefrihet. [52]

- **Challenge – response**
  Tekniken innebär att innan mottagande av ett e-postmeddelande krävs avsändaren på ett manuellt svar, som bekräftar att meddelandet inte är spam och att avsändaren är en människa. Enligt företaget Earthlinks vice VD Jim Anderson så finns det inget som indikerar att spammare i dagsläget hanterar den tekniken. Flera experter pekar dock på problemet att trafikmängden för e-post kan dubbleras, samt att den är problematisk för stora implementationer. [51]

- **Effektivare filter**
  I dagsläget är spamfilterutvecklingen en ständig kamp mot spammare. Spammare uppfinner ständigt nya metoder för att slinka förbi de mest sofistikerade filtren. Enligt IBMs forsknings- och utvecklingsgrupp X-Force var 40 % av all spam bildbaserad under slutet av 2006. Det finns flera olika sätt att lura de än så länge underlägsna bildspamfilter som finns att tillgå, genom att använda sig av bland annat animerade bilder eller bilder med flera eller transparent bakgrund. Gemensamt för dem alla är att det är mycket svårt att maskinellt avgöra om det rör sig om spam, samt att det är en mycket resursintensiv process att avkoda och tolka innehållet i bilden. [52, 53]

- **Global opt-out-lista**
  Möter problemet med att få till stånd en internationell lagstiftning och att kunna se till att den efterlevs. [52]

6. Slutsatser

Våra försök visar att SpamAssassin i sin grundkonfiguration inte uppnår hundraprocentig träffsäkerhet. Därmed finns det risk för falska positiva eller falska negativa resultat och är något att beakta om man väljer att med automatik kasta e-post som nått tröskelvärdet.

Försöken visar även på vikten av tillgång till de publika blocklistorna för en korrekt bedömning av inkommande e-post. Vår bedömning är att blocklistorna är det viktigaste verktyget som SpamAssassin har att tillgå om det Bayesiska filtret ej är upplärt.

Vi tror att vi skulle fått ett mer tillförlitligt resultat om vi lärt upp det Bayesiska filtret, men då vi saknade tillräcklig mängd spam och legitim e-post var detta ej möjligt.

På grund av att vi inte använde programvara för bildavläsning uppnåddes mindre bra resultat i avseende filtrering av bildbaserad spam.

6.1. Slutsats från virtuell miljö

Vår bedömning av anledningen till SpamAssassins låga träffsäkerhet främst beträffande bildspam i den simulerade miljön är följande:

- De datorer som på Internet skickar spam registreras i olika blocklistor. Eftersom all e-post som skickades kom från samma IP-adress är detta ett kriterium som Spamassassin ej kunnat beakta vid test i den virtuella miljön.

- Tekniken för filtrering av bildbaserad spam är i dagsläget långt efter tekniken för dess textbaserade motsvarighet. [52, 53] SpamAssassin i sitt grundutförande har ej stöd för OCR-läsning av bilder bifogade i e-postmeddelanden, men kan utökas med tilläggsprogramvara.

- På grund av otillräcklig mängd spam och legitim e-post har vi ej kunnat lära upp de inbyggda filterfunktionerna i SpamAssassin. Därmed har SpamAssassin fått arbeta efter de grundregler som programmet levereras med.

6.2. Slutsats från praktiskt försök

Det måste tas i beaktande att försöksperioden var tämligen kort, samt att domänen som användes ej sedan tidigare var etablerad på Internet. Därmed var domänen och dess mottagaradresser varken kända eller utspridda i spamkretsar. Det innebar att mängden mottagen e-post inte nådde upp till den proportion spam respektive legitim e-post som enligt våra källor för närvarande råder på Internet.

Urvalet av spam kan inte ses som representativt för all den spam som cirkulerar på Internet, och därmed är det svårt att dra någon reell slutsats om SpamAssassins effektivitet.

Den andel legitim e-post som skickades till domänen var i hög majoritet nyhetsbrev som beställts av projektmedlemmarna, i syfte att cirkulera legitim e-post till domänen. Det skedde mycket lite mänsklig korrespondens till och från domänen, och därmed utgjordes så gott som all legitim e-post av opersonliga nyhetsbrev. SpamAssassin fick därmed aldrig tillfälle att analysera personlig korrespondens. Dock måste det nämnas att vid de tillfällen vi skickat e-post mellan gruppmedlemmarna har SpamAssassin inte bedömt breven komma ens i närheten av det förinställda tröskelvärdet 5.0.

7. Referenser

1. Justitiedepartementet (2006). *Offentlighetsprincipen*
http://www.sweden.gov.se/sb/d/1487/a/12528 [2007-05-04] (Elektronisk)

2. Anders Tandersten (2004). *Spam: Den nya tidens onlinegissel.* Magisteruppsats, Bibliotekshögskolan i Borås [2007-05-04] Även tillgänglig som:
http://dspace.bib.hb.se/dspace/bitstream/2320/1155/1/04-123.pdf

3. Commtouch (2007). *2007 Q1 Spam Trends: Botnets Continue Sending Devious Spam.*
http://www.commtouch.com/documents/Commtouch_2007_Q1_Spam_Trends.pdf [2007-05-04] (Elektronisk)

4. Postini (2007). *2007 Postini Communications Intelligence Report.*
http://www.postini.com/whitepapers/WP43-2007_CIR.pdf [2007-05-04] (Elektronisk)

5. Marshall Brain (2007). *How E-mail Works*
http://computer.howstuffworks.com/email.htm [2007-05-15] (Elektronisk)

6. Okänd (2003). *Q10001: How does email work?*
http://www.mailbox.net.uk/page.php?cid=44 [2007-05-15] (Elektronisk)

7. P.Resnick (2001). *RFC2882 Internet Message Format.*
http://tools.ietf.org/html/rfc2822 [2007-05-15] (Elektronisk)

8. Anita Urgell (2005). *Du har 923 nya mail…* Magisteruppsats, Handelshögskolan vid Göteborgs Universitet. [2007-05-15] Även tillgänglig som:
http://www.handels.gu.se/epc/archive/00004093/01/Nr%5F5%5FAU.pdf

9. Spamhaus (okänt). *The Definition of Spam*
http://www.spamhaus.org/definition.html [2007-05-15] (Elektronisk)

10. Nationalencyklopedin (2007-05-15). *Nationalencyklopedins Internettjänst NE.se*
http://www.ne.se/jsp/search/article.jsp?i_art_id=492000&i_word=spam [2007-05-15] (Elektronisk)

11. Scott Hazen Mueller (okänt). *What is Spam?*
http://spam.abuse.net/overview/whatisspam.shtml [2007-05-15] (Elektronisk)

12. Kaspersky Lab (okänt). *Spam – What exactly is it?*
http://www.viruslist.com/en/spam/info?chapter=153350526 [2007-05-15] (Elektronisk)

13. Spamhaus (2007-05-15). *Spamhaus TOP 10 Spam Origin Countries*
http://www.spamhaus.org/statistics/countries.lasso [2007-05-15] (Elektronisk)

14. Kaspersky Labs (okänt). *Contemporary Spammer Technologies*
http://www.viruslist.com/en/spam/info?chapter=153350528 [2007-05-15] (Elektronisk)

15. Safe-Mail.Net (okänt). *The Economy Of SPAM*
http://www.safe-mail.net/docs/SpamEconomy.html [2007-05-15] (Elektronisk)

16. Computer Sweden (2004-03-04). *Telia stänger spam-port*
http://www.idg.se/2.1085/1.18843 [2007-05-15] (Elektronisk)

17. Urban Lindstedt (2003-12-04). *BBB blockerar kundernas mejlservrar*
http://www.idg.se/2.1085/1.53437 [2007-05-15] (Elektronisk)

18. Gate Comm Software (Okänt). *Port 25 Blocking*
http://www.postcastserver.com/help/Port_25_Blocking.aspx [2007-05-21] (Elektronisk)

19. Konsumentverket (2004-03-29). *E-postreklam – spam*
http://www.epostreklam.konsumentverket.se [2007-05-15] (Elektronisk)

20. Konsumentverket (2007-02-12). *Regler om obeställd e-postreklam*
http://www.epostreklam.konsumentverket.se/mallar/sv/artikel.asp?lngCategoryId=1586&lngArticleId=3607 [2007-05-15] (Elektronisk)

21. Federal Trade Commission (2004-04). *The CAN-SPAM Act: Requirements for Commercial Emailers* http://www.ftc.gov/bcp/conline/pubs/buspubs/canspam.shtm [2007-05-16] (Elektronisk)

22. David McGuire (2003-12-17). *New Law Won't Can Spam, Critics Say*
http://www.washingtonpost.com/wp-dyn/articles/A5943-2003Dec16.html [2007-05-16] (Elektronisk)

23. Justin Mason (2004-06-02). *SpamAssassin Prehistory: filter.plx*
http://spamassassin.apache.org/prehistory/ [2007-04-10] (Elektronisk)

24. Justin Mason (2006-12-06). *SpamAssassin – Spamassassin Wiki*
http://wiki.apache.org/spamassassin/SpamAssassin [2007-05-16] (Elektronisk)

25. Okänd (okänt). *Test Performed: v3.1.x*
http://spamassassin.apache.org/tests_3_1_x.html [2007-05-15] (Elektronisk)

26. Justin Mason (2005-06-03). *SpamAssassin Rules – Spamassassin Wiki*
http://wiki.apache.org/spamassassin/SpamAssassinRules [2007-04-16] (Elektronisk)

27. Okänd (okänt). *Who was the Reverend Thomas Bayes?*
http://www.bayesian.org/bayesian/bayes.html [2007-05-16] (Elektronisk)

28. Okänd (okänt). *What is FuzzyOCR?*
http://fuzzyocr.own-hero.net/wiki/WhatisFuzzyOcr [2007-05-16] (Elektronisk)

29. Ola Sigurdsson (2007-04). *Bildskräp tvingar fram nya skydd.*
TechWorld #4 2007. (Tidskrift)

30. Okänd (okänt). *The Postfix Home page*
http://www.postfix.org [2007-05-16] (Elektronisk)

31. Okänd (okänt). *Postfix feature overview*
http://www.postfix.org/features.html [2007-05-16] (Elektronisk)

32. Ivar Abrahamsen (2006-11-27). *How to set up a mail server on a GNU / Linux system*
http://flurdy.com/docs/postfix/#conf_mta [2007-05-16] (Elektronisk)

33. Advosys Consulting Inc. (2006-01-02). *Fighting malware and spam with Postfix*
http://advosys.ca/papers/postfix-filtering.html [2007-05-16] (Elektronisk)

34. P. Hoffman, S. Harris (2006-09). *The Tao of IETF: A Novice's Guide to the Internet Engineering Task Force.*
http://www.ietf.org/tao.html [2007-05-16] (Elektronisk)

35. Jonathan B. Postel (1982-08). Simple Mail Transfer Protocol
http://www.ietf.org/rfc/rfc0821.txt [2007-05-21] (Elektronisk)

36. J. Klensin (2001-04). *Simple Mail Transfer Protocol*
http://tools.ietf.org/html/rfc2821 [2007-05-16] (Elektronisk)

37. Evan Harris (2003-08-21). *The Next Step in the Spam Control War: Greylisting*
http://projects.puremagic.com/greylisting/whitepaper.html [2007-05-16] (Elektronisk)

38. Jan Melin (2005-03-02). *Spam är inte längre något problem?*
http://nyteknik.se/art/39257 [2007-05-16] (Elektronisk)

39. Bjarne Lundgren (2004). *Greylisting*
http://www.greylisting.org/ [2007-05-16] (Elektronisk)

40. Okänd (okänt). *Distributed Checksum Clearinghouse*
http://www.rhyolite.com/anti-spam/dcc/ [2007-05-16] (Elektronisk)

41. Vipul Ved Prakash (2006-05-31). *Vipul's Razor v2 README*
http://razor.sourceforge.net/docs/doc.php?type=text&name=README [2007-05-16] (Elektronisk)

42. Okänd (okänt). *Steps to block spam with Pyzor*
http://pyzor.sourceforge.net/blocking.html [2007-05-16] (Elektronisk)

43. Okänd (okänt). *Mail::SpamAssassin::Plugin::DCC – perform DCC check of messages*
http://spamassassin.apache.org/full/3.2.x/doc/Mail_SpamAssassin_Plugin_DCC.html [2007-05-23] (Elektronisk)

44. Okänd (okänt). *Mail::SpamAssassin::Plugin::Pyzor – perform Pyzor check of messages*
http://spamassassin.apache.org/full/3.2.x/doc/Mail_SpamAssassin_Plugin_Pyzor.html [2003-05-23] (Elektronisk)

45. Okänd (okänt). *Mail::SpamAssassin::Plugin::Razor2 – perform Razor check of messages*
http://spamassassin.apache.org/full/3.2.x/doc/Mail_SpamAssassin_Plugin_Razor2.html [2007-05-23] (Elektronisk)

46. Spamhaus (okänt). *About Spamhaus*
http://www.spamhaus.org/organization/index.lasso [2007-05-16] (Elektronisk)

47. Spamhaus (okänt). *Understanding DNSBL Filtering*
http://www.spamhaus.org/dnsbl_function.html [2007-05-16] (Elektronisk)

48. Spamhaus (okänt). *SBL Policy & Listing Criteria*
http://www.spamhaus.org/sbl/policy.html [2007-05-16] (Elektronisk)

49. Spamhaus (okänt). *Frequently Asked Questions (FAQ)*
http://www.spamhaus.org/faq/answers.lasso?section=Glossary#169 [2007-05-16] (Elektronisk)

50. Okänd (2005-05-04). *TestingInstallation*
http://wiki.apache.org/spamassassin/TestingInstallation [2007-05-16] (Elektronisk)

51. Jane Weaver (2006-07-10). *How to end spam in the future*
http://www.msnbc.msn.com/id/3078599/ [2007-05-16] (Elektronisk)

52. Brad Templeton (okänt). *Reflections on the 25th Anniversary of Spam*
http://www.templetons.com/brad/spam/spam25.html [2007-05-16] (Elektronisk)

53. Carsten Dietrich (2007-03-20). *The future of image spam*
http://scmagazine.com/us/news/article/645023/the-future-image-spam [2007-05-16] (Elektronisk)