COMBATE AO SPAM – UMA ABORDAGEM RADICAL

José Roberto Bollis Gimenez
UNG / UNESP

jroberto@unesp.br
(endereço restrito)

Roteiro da Palestra

O Problema do SPAM

Alternativas possíveis para combater o SPAM

Alternativas Técnicas

- Técnicas de bloqueio por origem
- Técnicas de bloqueio por conteúdo
- Experiência do apresentador

Meios invasivos de propaganda

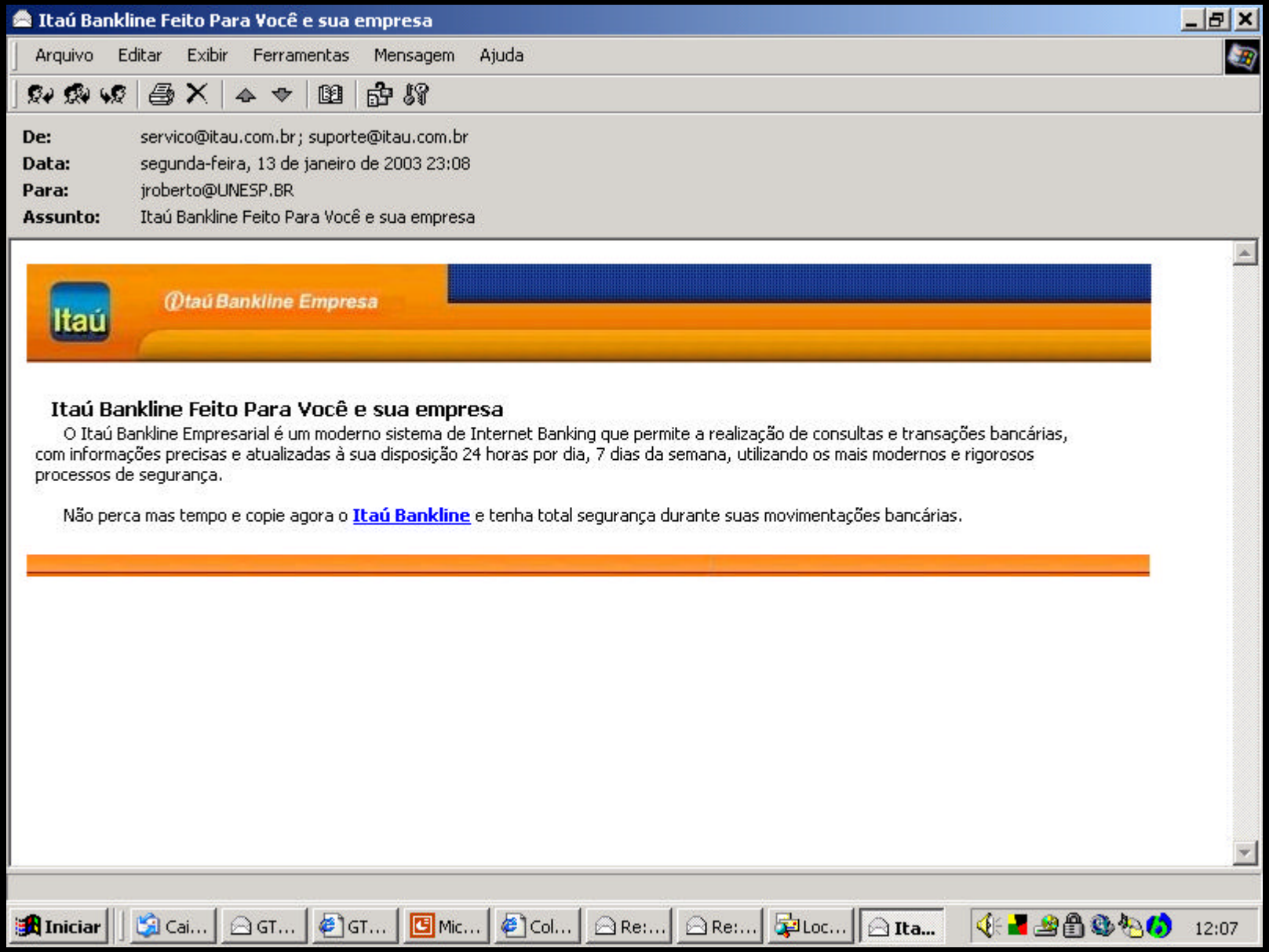
- Telemarketing
- Correio eletrônico (SPAM)
- Correio convencional (mala direta)
- Moças entregando anúncios em semáforos
- Comerciais indesejáveis em canais pagos de televisão
- Merchandising no meio de novelas
- Etc

Por que o SPAM incomoda?

- Porque ele cresce exageradamente
- Porque ele ameaça a existência do correio eletrônico
- Porque o material propagado é de péssima qualidade

Quem são os spammers ?

- Gente da pior espécie
  - Falsificação de headers
  - Uso de relays indevidos
  - Uso de anonimato
  - Declarações mentirosas
- Gente da melhor espécie
  - Empresas “sérias” também fazem uso do SPAM


Itaú Bankline Feito Para Você e sua empresa
Arquivo Editar Exibir Ferramentas Mensagem Ajuda
De: servico@itau.com.br; suporte@itau.com.br
Data: segunda-feira, 13 de janeiro de 2003 23:08
Para: jroberto@UNESP.BR
Assunto: Itaú Bankline Feito Para Você e sua empresa
Itaú
Itaú Bankline Empresa
Itaú Bankline Feito Para Você e sua empresa
O Itaú Bankline Empresarial é um moderno sistema de Internet Banking que permite a realização de consultas e transações bancárias, com informações precisas e atualizadas à sua disposição 24 horas por dia, 7 dias da semana, utilizando os mais modernos e rigorosos processos de segurança.
Não perca mas tempo e copie agora o Itaú Bankline e tenha total segurança durante suas movimentações bancárias.
Iniciar
Cai...
GT...
GT...
Mic...
Col...
Re:...
Re:...
Loc...
Ita...
12:07

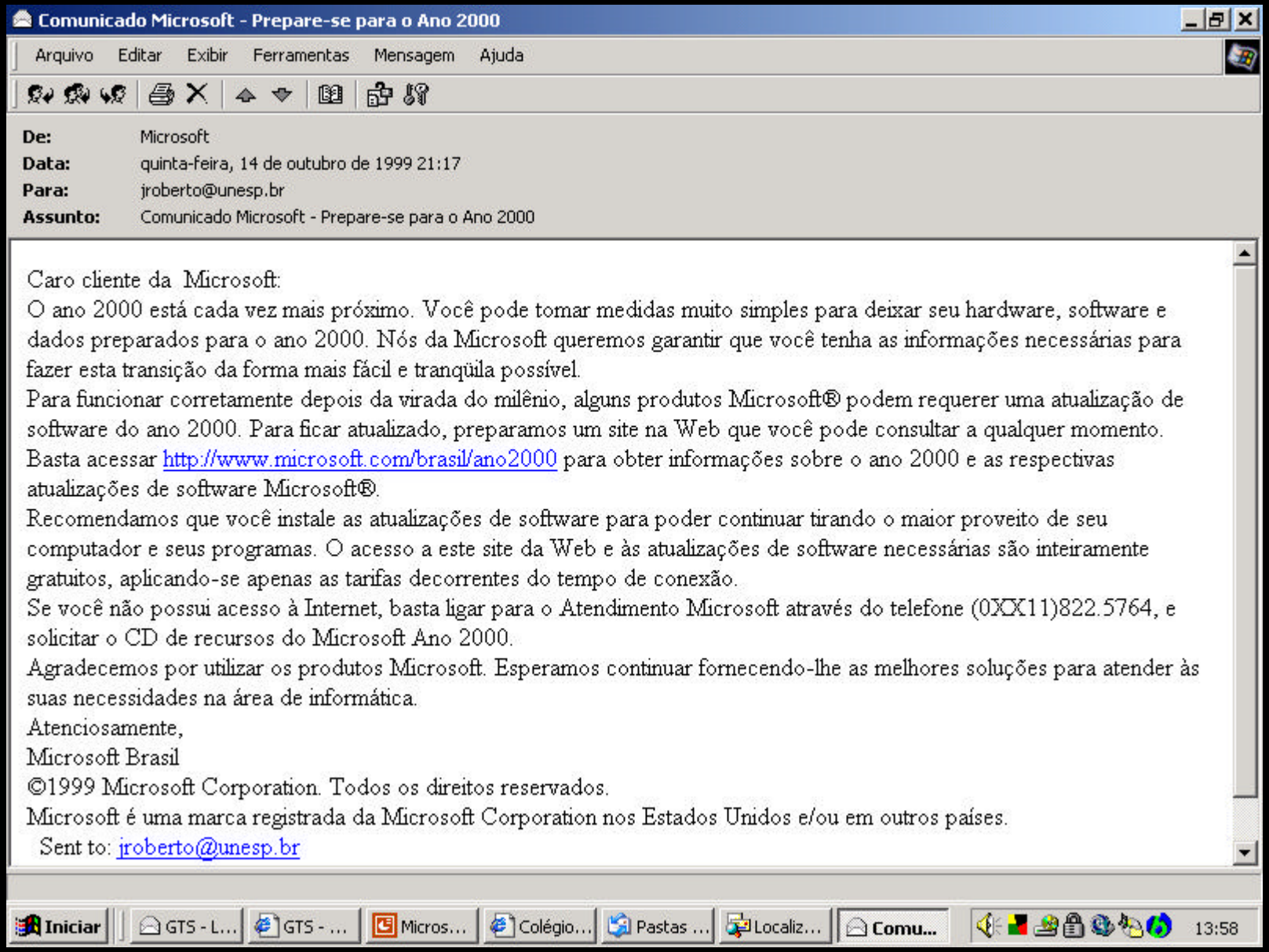
Fw: O futuro em seu e-mail

Arquivo Editar Exibir Ferramentas Mensagem Ajuda

De: Jose Roberto B. Gimenez
Data: sexta-feira, 29 de novembro de 2002 07:44
Para: investor@nortelnetworks.com
Assunto: Fw: O futuro em seu e-mail

----- Original Message -----
From: Nortel Networks
To: jroberto@unesp.br
Sent: Thursday, November 28, 2002 9:42 PM
Subject: O futuro em seu e-mail

Iniciar | GT... | GT... | Mic... | Col... | Cai... | Co... | Loc... | Fw: ... | 14:13


Comunicado Microsoft - Prepare-se para o Ano 2000
Arquivo Editar Exibir Ferramentas Mensagem Ajuda
De: Microsoft
Data: quinta-feira, 14 de outubro de 1999 21:17
Para: jroberto@unesp.br
Assunto: Comunicado Microsoft - Prepare-se para o Ano 2000
Caro cliente da Microsoft:
O ano 2000 está cada vez mais próximo. Você pode tomar medidas muito simples para deixar seu hardware, software e dados preparados para o ano 2000. Nós da Microsoft queremos garantir que você tenha as informações necessárias para fazer esta transição da forma mais fácil e tranqüila possível.
Para funcionar corretamente depois da virada do milênio, alguns produtos Microsoft® podem requerer uma atualização de software do ano 2000. Para ficar atualizado, preparamos um site na Web que você pode consultar a qualquer momento.
Basta acessar http://www.microsoft.com/brasil/ano2000 para obter informações sobre o ano 2000 e as respectivas atualizações de software Microsoft®.
Recomendamos que você instale as atualizações de software para poder continuar tirando o maior proveito de seu computador e seus programas. O acesso a este site da Web e às atualizações de software necessárias são inteiramente gratuitos, aplicando-se apenas as tarifas decorrentes do tempo de conexão.
Se você não possui acesso à Internet, basta ligar para o Atendimento Microsoft através do telefone (0XX11)822.5764, e solicitar o CD de recursos do Microsoft Ano 2000.
Agradecemos por utilizar os produtos Microsoft. Esperamos continuar fornecendo-lhe as melhores soluções para atender às suas necessidades na área de informática.
Atenciosamente,
Microsoft Brasil
©1999 Microsoft Corporation. Todos os direitos reservados.
Microsoft é uma marca registrada da Microsoft Corporation nos Estados Unidos e/ou em outros países.
Sent to: jroberto@unesp.br
Iniciar
GTS - L...
GTS - ...
Micros...
Colégio...
Pastas ...
Localiz...
Comu...
13:58

Abordagens para combater o SPAM

- Legislação apropriada que proíba o SPAM
- Enforcement adequado
- Sobretaxa do serviço
- Conscientização do usuário
- Uso de Recursos Técnicos

Legislação

- A legislação é limitada a um determinado país, enquanto Internet não tem fronteiras.
- Em mais de 30 anos de existência da rede, não se conhecem leis apropriadas nesta área.
- O congresso não consegue resolver problemas básicos como: fome, criminalidade, distribuição de renda, etc,
- Em certas áreas temos leis excelentes, porém não resolvem nada.

Ações na Justiça / Enforcement

- Justiça é feita por advogados e juízes.
- Advogados trabalham para a parte que mais paga.
- Os casos conhecidos (no Brasil) de ação contra spammers são desastrosos.
- Os juízes confundem SPAM com Telemarketing (e o pior, é que estão certos).

Fluxograma de uma condenação

begin;

if (Crime prescreveu), then: go to **FREE**;

else if (Réu primário), then: go to **FREE;**

else if (Não foi fragrante), then: go to **FREE**;

else if (Pagou fiança), then: go to **FREE**;

else if (Habeas Corpus), then, go to **FREE**;

else if (Nível universitário); then: go to **CELA_ESPECIAL**;

if (Não tem $$ para pagar advogado); then: go to **CANA**;

end.

Sobretaxar o Correio Eletrônico

- Poderia estragar a melhor coisa que o sistema tem – a gratuidade.
- Quem ficaria com o dinheiro das taxas?
- Em que moeda seria pago?
- Os spammers seriam mesmo controlados?
- Esta medida não estaria legalizando o SPAM?
- O interesse em arrecadar com a taxa não acabaria incentivando o SPAM?

Conscientizar o usuário

É comum ouvir dos usuários expressões como:

- Fico tão triste quando encontro minha caixa de mensagens vazia.
- É tão agradável perceber que alguém se lembrou de mim.
- Se você não gosta de receber estas mensagens, basta pedir, que eles retiram você da lista.

Solução Técnica

- Quando encontrada é rápida, eficiente e não depende de discussões em assembléias.
- O técnico sempre procura resolver o problema, e não tirar proveito da situação.
- O técnico entende bem o problema e não confunde PVC com Policloreto de Vinila.

Quem deve combater o SPAM?

ADMINISTRADORES

- Possuem maior poder de ação.
- Podem implantar soluções unificadas.
- E depois...., têm que sofrer as conseqüências.

USUÁRIOS

- Conseguem diferenciar melhor um SPAM de uma mensagem legítima.
- Podem aplicar a sua própria regra.

Dificuldades em combater tecnicamente o SPAM

- Os headers normalmente são falsos
- A origem é falsa
- A pessoa é falsa
- O produto vendido é falso
- Os disclaimers são falsos

INFORMAÇÕES PELO SITE:
www.institutouniversalalpha.com.br

OBS: Esta mensagem não é um spam, visto que somente estará sendo enviado uma única vez, e também contém uma forma de ser removida, é um e-mail normal como tantos outros que você recebe, não estamos invadindo sua privacidade e enviar um e-mail não é crime, desde que não contenha mensagens que possam causar danos ao usuário. Caso queria remover seu endereço de nossa lista, basta enviar um e-mail para excluame@institutouniversalalpha.com.br, que seu e-mail será removido de nossa lista definitivamente. Desculpe-nos caso tenhamos lhe importunado com nosso e-mail de divulgação.
Obrigado!

•O Rei do e-mail.
Compre com quem é pioneiro em mala direta digital.

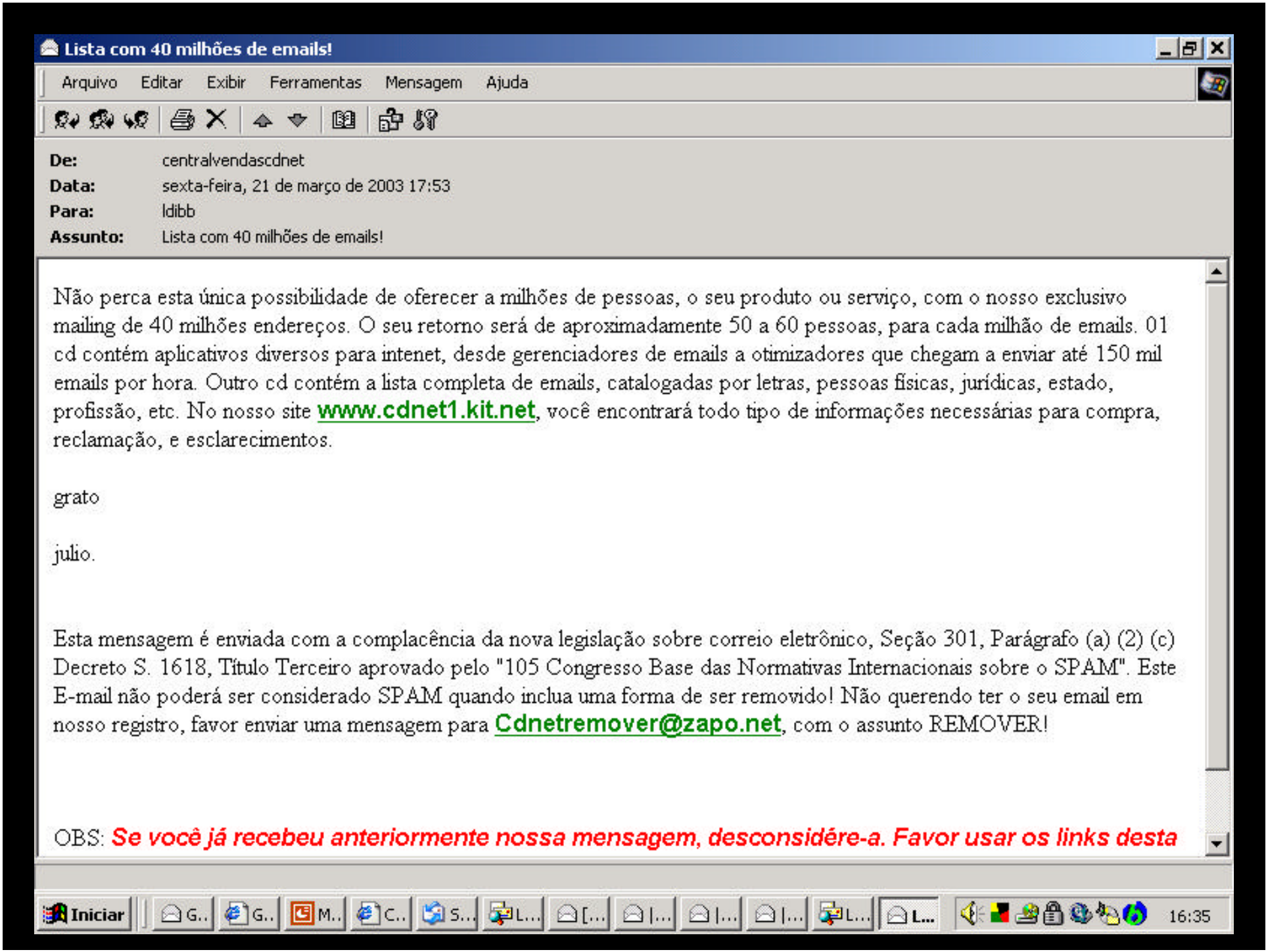
•ADQUIRA JÁ !!!
FAÇA SEU PEDIDO PELO SITE
www.oreidoemail.kit.net

OU LIGUE-NOS (48) 9112 9279

"E-mail Legal"
Em conformidade com o Projeto
de Lei "Anti-Spam"

Mensagem Eletrônica de Publicidade de Produtos e Serviços. A nova etiqueta da internet. Esse e-mail tem conteúdo institucional e pode ser facilmente filtrado pelo seu software de e-mail.

Caso não queira receber mais
as nossas mensagens
publicitárias clique aqui!


Lista com 40 milhões de emails!
Arquivo Editar Exibir Ferramentas Mensagem Ajuda
De: centralvendascdnet
Data: sexta-feira, 21 de março de 2003 17:53
Para: ldibb
Assunto: Lista com 40 milhões de emails!
Não perca esta única possibilidade de oferecer a milhões de pessoas, o seu produto ou serviço, com o nosso exclusivo mailing de 40 milhões endereços. O seu retorno será de aproximadamente 50 a 60 pessoas, para cada milhão de emails. 01 cd contém aplicativos diversos para intenet, desde gerenciadores de emails a otimizadores que chegam a enviar até 150 mil emails por hora. Outro cd contém a lista completa de emails, catalogadas por letras, pessoas físicas, jurídicas, estado, profissão, etc. No nosso site www.cdnet1.kit.net, você encontrará todo tipo de informações necessárias para compra, reclamação, e esclarecimentos.
grato
julio.
Esta mensagem é enviada com a complacência da nova legislação sobre correio eletrônico, Seção 301, Parágrafo (a) (2) (c) Decreto S. 1618, Título Terceiro aprovado pelo "105 Congresso Base das Normativas Internacionais sobre o SPAM". Este E-mail não poderá ser considerado SPAM quando inclua uma forma de ser removido! Não querendo ter o seu email em nosso registro, favor enviar uma mensagem para Cdnetremover@zapo.net, com o assunto REMOVER!
OBS: Se você já recebeu anteriormente nossa mensagem, desconsidére-a. Favor usar os links desta
Iniciar
16:35

Meios para combater tecnicamente o SPAM

- Bloqueio por endereço de destino
  - RBL
  - Teergrubing
  - TMDA
- Bloqueio por análise de conteúdo
  - SpamAssassin
  - Bogofilter (Filtragem Bayesiana)

Utilização de RBL

- Serviço mantido e arbitrado pelo administrador.
- Presume a inexistência de Relays abertos.
- Não funciona para bloquear endereços individuais. Deve-se bloquear toda a rede de origem.

RBL - Realtime Blackhole List

http://mail-abuse.org - MAPS

Problemas do RBL

- Mensagens legítimas também provêm de redes que são coniventes com o abuso.
- Poucos provedores cedem, ou mesmo respondem, às queixas de SPAM.
- O administrador acaba sendo crucificado quando uma mensagem legítima é bloqueada.

Teergrubing Wrapper

- Serviço mantido e arbitrado pelo administrador.
- Trata de forma “especial” as conexões vindas de determinados domínios.
- Não impede o recebimento de mensagens legítimas dos domínios listados.

Problemas do Teergrubing

- Os mesmos do RBL, minimizados por não impedir totalmente o recebimento de emails legítimos.
- É necessária a disseminação do sistema. Um único programa não produz resultados.

SpamAssassin

- Bloqueio pela análise de características que são próprias de SPAM.
- A ocorrência de tais características é pontuada e ocorre o bloqueio quando um threshold é atingido.

http://spamassassin.org

SPAM: ---------- Start SpamAssassin results ----------
SPAM:
SPAM: Content analysis details: (17.0 hits, 8 required)
SPAM: Hit! (1.3 points) From: termina com números
SPAM: Hit! (3.5 points) Priority: mensagem enviada com alta prioridade
SPAM: Hit! (2.0 points) Invalid Date: header (no timezone)
SPAM: Hit! (1.0 point) BODY: /responda indicando no assunto/i
SPAM: Hit! (1.0 point) BODY: /REMOVER/
SPAM: Hit! (1.0 point) BODY: /é enviada com a complacência/i
SPAM: Hit! (1.0 point) BODY: /S. 1618/i
SPAM: Hit! (1.0 point) BODY: /não poderá ser considerado SPAM/i
SPAM: Hit! (0.7 points) BODY: Contém linha com mais de 199 caracteres
SPAM: Hit! (5.5 points) Recebido através de IP bloqueado em rbl.unesp.br
SPAM: [RBL check: found 183.27.204.200.rbl.unesp.br.]
SPAM:
SPAM: ---------- End of SpamAssassin results ----------

Problemas do SpamAssassin

- Os Spammers começam adaptar seus textos para driblar a filtragem.

Exemplo:

Subject: D I V U L G A Ç Ã O

Body: F.R.E.E

Filtragem Bayesiana

Teorema de Bayes

$$P(A|B) \ ? \ \frac{P(A)}{P(B)} P(B|A)$$

A = A mensagem recebida ser um SPAM

B = A mensagem recebida ter determinada característica

Probabilidade combinada

$$P(A \mid B_1, B_2, .., B_N) \; ? \; \frac{? \; \prod_{i?1}^{N} P(B_i \mid A)}{? \; \prod_{i?1}^{N} P(B_i \mid A) \; ? \; ? \; \prod_{i?1}^{N} [1 ? \; P(B_i \mid A)]}$$

A = A mensagem recebida ser um SPAM

B_1 = A mensagem recebida ter a característica 1

B_2 = A mensagem recebida ter a característica 2

.....

B_N = A mensagem recebida ter a característica N

Filtragem Bayesiana

- É criado um arquivo com um grande número de SPAMs (milhares).
- Cada palavra recebe um valor probabilístico, conforme sua freqüência no texto do SPAM.
- A Fórmula da Probabilidade Combinada é aplicada.
- Se um determinado valor é ultrapassado, a mensagem é considerada SPAM.

Bloqueio pelo conteúdo – Desvantagens

- Ocorrência de falsos negativos - alguns SPAMs continuam incomodando.
- Ocorrência de falsos positivos - todos os SPAMs continuam incomodando (procurar no lixo do mailbox).

Outras desvantagens

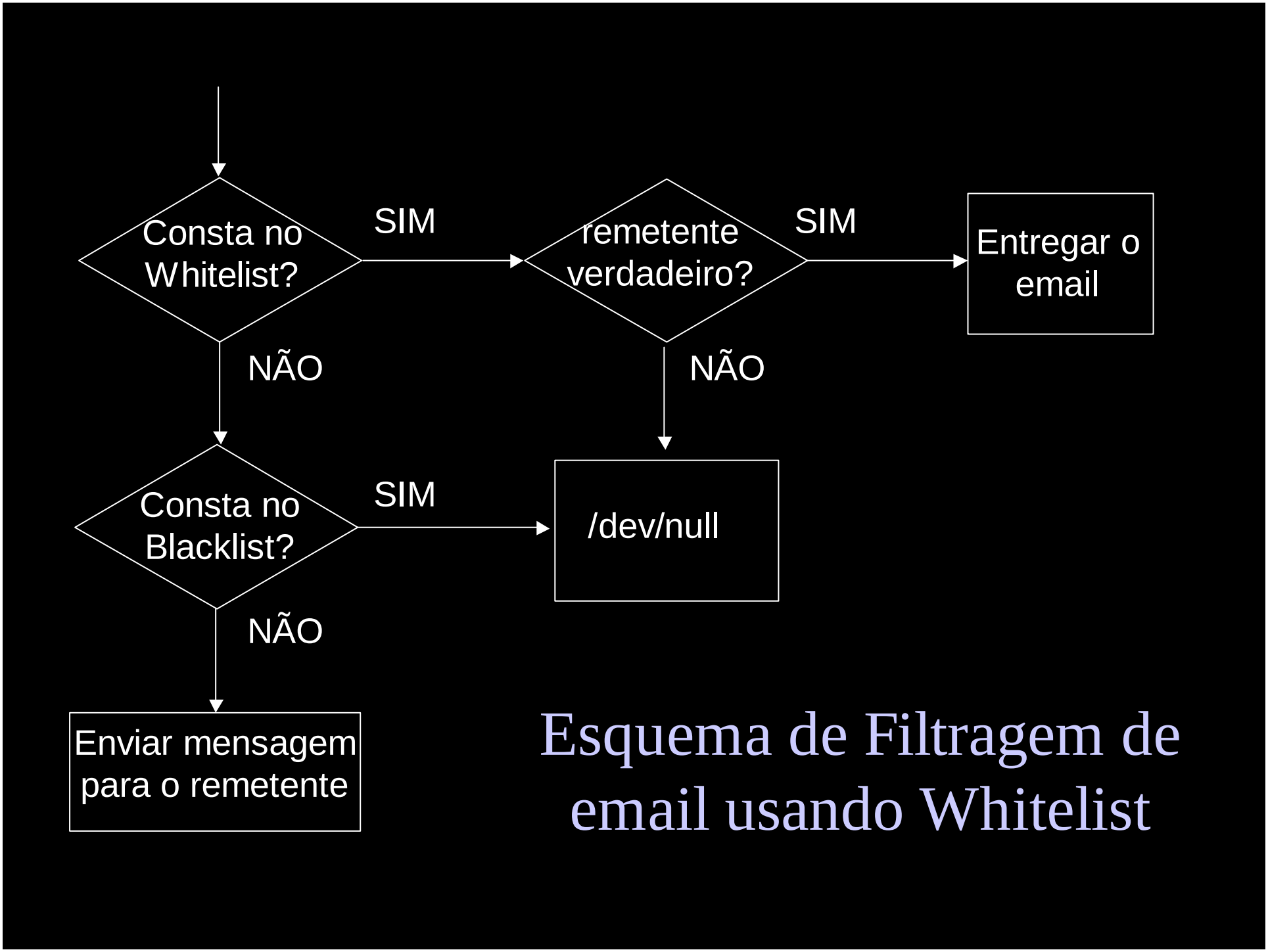
- Os SPAMs começam a ficar parecidos com as mensagens legítimas, dificultando a filtragem automática.
- Eles ficarão tão mais parecidos quanto maior for a vontade dos spammers que você os receba.

Whitelist - Solução Definitiva

- Controle total sobre quem pode enviar mensagens.
- Não existe o perigo do falso negativo (pelo menos em tese).
- Não existe o perigo do falso positivo (para o pessoal do whitelist).

Implementação do Whitelist

- O header do email é analisado e comparado com uma lista de remetentes permitidos (PROCMAIL).
- Uma mensagem de aviso é enviada para o remetente no caso deste não constar da lista (FORMAIL).
- Deve-se também chegar a autenticidade da origem.



Esquema de Filtragem de email usando Whitelist

Desvantagem do Whitelist

- Implementação no servidor UNIX, usando PROCMAIL e FORMAIL (nada próprio para usuários comuns).
- Endereços de Abuse, Postmaster, SOA, Ouvidoria, SAC, etc, não devem usar esta abordagem.
- Os Spammer começam tentar adivinhar o conteúdo de sua lista.

Outras desvantagens...

- Uma pessoa que não te conhece e que esteja querendo “divulgar” um produto interessante, não conseguirá fazê-lo...

Que bom!!!

Links Interessantes

Teergrubing Wrapper

http://www.iks-jena.de/mitarb/lutz/usenet/antispam.html

Filtragem bayesiana

http://www.paulgraham.com/spam.html

Probabilidade Combinada

http://www.mathpages.com/home/kmath267.htm

Produtos que utilizam Filtragem Bayesiana

http://bogofilter.sourceforge.net/

http://www.mozilla.org/mailnews/spam.html

Links Interessantes

Produtos que utilizam Filtragem por origem

- http://www.tmda.net/
- http://www.mailabuse.org

Produtos que utilizam Filtragem por conteúdo

- http://spamassassin.org/tag/