USO DE *GREYLISTS* E CONTROLE DE *RELAYS* ABERTOS COMO TÉCNICA DE BLOQUEIO DE SPAM: TRANSFERINDO A RESPONSABILIDADE PARA O REMETENTE

Samuel Pereira Dias
Universidade Federal de Lavras - Curso de Pós-Graduação Lato Sensu Administração em Redes Linux
Caixa Postal 3037 – CEP 37200-000 – Lavras-MG
samuel@ginux.ufla.br

Joaquim Quinteiro Uchôa
Universidade Federal de Lavras - Curso de Pós-Graduação Lato Sensu Administração em Redes Linux
Caixa Postal 3037 – CEP 37200-000 – Lavras-MG
joukim@ginux.ufla.br

RESUMO

O recebimento de mensagens indesejadas, conhecido como *spam*, é um dos problemas mais sérios enfrentados pelos administradores de rede atualmente. Em geral, as técnicas tradicionais de filtragem consomem grande esforço computacional para separar mensagens legítimas do *spam* e, portanto, apresentam problemas de desempenho em máquinas mais modestas. O presente trabalho apresenta o caso de uso da técnica de *greylists* aliada à verificação de servidores com *relay* aberto para esse controle. Os resultados obtidos mostraram-se favoráveis à adoção da técnica por sua simplicidade e baixo consumo de recursos.

PALAVRAS-CHAVE

E-mail, spam, greylist, open relay

1. INTRODUÇÃO

Atualmente, o bloqueio de mensagens comerciais não solicitadas, também conhecidas como *spam* constitui um dos maiores problemas dos administradores de redes. O consumo de banda de rede e espaço de armazenamento, a insatisfação crescente dos usuários são variáveis que ocupam a mente do administrador e transfornam o sistema de correio eletrônico em algo de gerenciamento complexo.

Segundo [CERT.BR, 2005], os usuários também são afetados de diferentes formas, tais como o não recebimento de mensagens por extrapolar a quota devido ao excesso de *spam,* perda de produtividade e gasto desnecessário de tempo analisando e separando mensagens indesejadas. Além disso, existem riscos de se envolver em esquemas fraudulentos que induzem o usuário a acessar páginas clonadas de instituições financeiras, provocando prejuízos (técnica conhecida como *phishing scam*). Independentemente da forma de acesso à *internet*, quem arca com os custos do envio de *spam* é o usuário, seja na mensalidade do provedor de acesso, no armazenamento, na filtragem ou na transmissão para seu micro pessoal.

Neste trabalho, é apresentada a solução adotada no Departamento de Ciência da Computação de uma instituição de ensino superior. O sistema de correio eletrônico abrange cerca de trezentos endereços válidos, entre alunos, professores e listas de discussão internas e externas. Nesse ambiente, os danos causados pelo *spam* na transmissão, armazenamento e processamento das mensagens exigiram medidas que minimizassem os impactos sem extrapolar a capacidade de processamento do servidor. Para apresentar a solução, na Seção 2 é feita a discussão dos métodos tradicionais para tratamento do problema, contraposto na Seção 3 pelo método alternativo selecionado pela administração da rede. Na Seção 4 encontra-se a análise dos resultados obtidos e na Seção 5, as conclusões deste trabalho e possíveis trabalhos futuros.

2. TÉCNICAS TRADICIONAIS DE CONTROLE DE SPAM

Nos últimos anos, muitas técnicas de filtragem e bloqueio de *spam* têm sido propostas e algumas delas são apresentadas nesta seção. Em geral, para cada mensagem é associado um *score* a partir da análise de seu conteúdo completo (cabeçalhos, assunto e corpo) e esse *score* indica o quanto essa mensagem parece ser um *spam* ou uma mensagem legítima. Por esse motivo, não são usadas como técnicas de bloqueio, mas de identificação da mensagem. O mecanismo de identificação não decide se uma mensagem será descartada ou entregue à caixa postal do usuário, uma vez que podem ocorrer mensagens legítimas identificadas como *spam* e mensagens indesejadas, como legítimas; respectivamente conhecidas como falsos positivos e falsos negativos.

Várias abordagens vêm sendo propostas para minimizar falsos positivos e aumentar a eficiência da filtragem. Pode-se citar como técnicas correntes os modelos Bayesianos para análise da propabilidade de cada palavra ser encontrada em uma mensagem legítima ou em um *spam* [SAHAMI et al., 1998], os métodos adaptativos baseados em algoritmos de *data mining* [MANCO et al., 2002], os métodos baseados em sistemas imunes artificiais [ODA, WHITE, 2003a; ODA, WHITE, 2003b]. Também merecem destaque os modelos de aprendizado de máquina para filtragem [TRETYAKOV, 2004], os métodos baseados em *boosting trees* [CARRERAS, MÁRQUEZ, 2001] e filtragem por aproximação baseada em casos [CUNNINGHAM et al., 2003]. Existe também a proposta de adoção de uma moeda virtual para troca de mensagens [TURNER, HAVEY, 2004], visando reduzir a quantidade simultânea de mensagens que cada servidor de *e-mail* é capaz de enviar.

Em geral, a literatura apresenta os modelos bayesianos como ponto de referência nas análises de eficiência, com valores superiores a 95% de acerto. Entretanto, com exceção da adoção de moeda virtual, a maior parte das técnicas requer processamento das mensagens no servidor do destinatário, consumindo recursos computacionais para a realização da filtragem. Em equipamentos menos robustos (com relação a processamento, memória e dispositivos de armazenamento), dependendo do número de mensagens recebidas, sua utilização não se torna atrativa, exigindo alternativas de menor custo computacional.

3. MÉTODOS ALTERNATIVOS PARA BLOQUEIO DE SPAM

As técnicas apresentadas na Seção 2 apresentam como inconveniente o processamento local para identificar as mensagens que podem ser *spam*. No entanto, muitas instituições não dispõem de um servidor de *e-mail* dedicado e com alto poder de processamento. Na maioria dos casos, são máquinas comuns que provêem não apenas o serviço de *e-mail*, mas também hospedam páginas *web*, compartilham arquivos dos usuários com as estações, entre outros serviços de rede. Nesse aspecto, apesar de minimizar os impactos do *spam*, o custo do processamento continua com o servidor do destinatário. Como o destinatário fica com os custos (financeiros e computacionais) de processar o *spam*, a função de *spammer* torna-se atrativa. De alguma forma, parte do custo de envio da mensagem deve ser partilhado pelo remetente, no caso, o custo de processamento, reduzindo o número de mensagens indesejadas enviadas em lote.

3.1. Mecanismo de Envio de Mensagens

O envio de correio eletrônico envolve diversos componentes, que possuem seu funcionamento detalhados na literatura, como por exemplo, em [NEMETH *et al.*, 2001]. Entre esses componentes, pode-se citar o cliente de *e-mail*, também conhecido como MUA (*Mail User Agent*), o servidor de envio, chamado MTA (*Mail Transfer Agent*), através do protocolo SMTP (*Simple Mail Transport Protocol*) [POSTEL, 1982]. Existem outros componentes, como o MDA (*Mail Delivery Agent*), que é usado para a entrega e acesso à caixa postal do usuário, como os protocolos POP3 (*Post Office Protocol*) [MYERS, ROSE, 1996] e IMAP4 (*Internet Message Access Protocol*) [CRISPIN, 1994].

O *spammer* aproveita-se da comunicação entre os MTAs. O protocolo SMTP é composto, em sua versão mais simples, por oito comandos que fazem a conexão e o envio da mensagem [STEVENS, 1994]. Não inclui nativamente qualquer mecanismo de autenticação de usuários. Existe uma extensão proposta por [MYERS, 1999] que permite o uso de mecanismos de autenticação, mas esta solução não é amplamente

adotada. Além disso, nos primeiros anos da *internet*, era comum permitir que vários MTAs fossem usados como intermediários na transmissão da mensagem, o que é chamado *open relay*, ou seja, especificar qual rota a mensagem deveria seguir até chegar em seu destinatário. A relação entre *open relays* e o envio de *spam* é apresentada na Seção 3.3.

3.2. O Uso de Greylisting como Método de Bloqueio

O protocolo SMTP [POSTEL, 1982] distingue duas categorias de erro: temporários (transitórios) e permanentes. Em um erro temporário, o MTA remetente pode enviar a qualquer tempo a mesma seqüência de comandos, visando a entrega da mensagem. A ocorrência de erro permanente sinaliza ao MTA remetente que a operação não pode ser repetida.

Os *spammers* usam MTAs modificados para permitir o envio de um grande número de mensagens no menor tempo possível. Para alcançar esse objetivo, esses MTAs não processam mensagens de erro do MTA destinatário. Essa característica pode ser explorada pelo administrador para evitar o recebimento de *spam* em seu domínio. Em uma seqüência normal de negociação SMTP para o envio da mensagem, o MTA remetente envia sua identificação (seu nome de domínio) e os endereços do remetente e do destinatário. Essas informações são chamadas envelope da mensagem[1]. De um modo geral, o *spammer* tende a fraudar sua identificação e o remetente, mas o MTA, devido à conexão TCP/IP utilizada, conhece o verdadeiro endereço de origem da conexão. Essas informações são utilizadas pelo sistema de *greylisting*.

Segundo [HARRIS, 2003], o funcionamento do *greylisting* enquadra-se no meio termo entre o *whitelisting* (endereços de onde se aceita incondicionalmente as mensagens) e *blacklisting* (endereços de onde devem ser rejeitadas as mensagens). O que é feito é uma postergação da entrega da mensagem, usando o próprio protocolo SMTP. Em seu funcionamento normal, após o envio do envelope da mensagem, o MTA destinatário está pronto para receber o conteúdo da mensagem. O funcionamento do *greylisting* está na interrupção do fluxo da transmissão nesse ponto: quando uma caixa postal recebe uma nova mensagem de um contato desconhecido (envelope não listado nas tabelas do sistema), a mensagem é temporariamente rejeitada. Isto ocorre na camada SMTP e é transparente para o usuário final. Em seguida, esse envelope é armazenado nas tabelas do banco de dados mantido pelo sistema de *greylisting*, com outras informações que permitem identificar sua origem real (endereço IP – *Internet Protocol*) e quando a mensagem foi recebida e recusada pela primeira vez. Em um curto intervalo de tempo, servidores de SMTP, que aderem corretamente aos padrões do protocolo, farão uma nova tentativa de envio. Ao reenviar, o sistema encontrará as informações da mensagem na base de dados, liberando sua entrega. Como o MTA do *spammer,* em geral, não processa mensagens de erro para reenviar a mensagem, uma nova tentativa de entrega não será realizada.

3.3. Recusando Mensagens de Open Relays

Neste ponto, pode ser notado que um *spammer* que utiliza um servidor com *open relay* pode burlar o mecanismo do *greylisting*, uma vez que o MTA explorado pode ter uma implementação que faça novas tentativas de entrega da mensagem e o protocolo SMTP padrão não inclui mecanismos de autenticação. Como uma mensagem de *spam* pode vir dessa fonte, torna-se necessário utilizar algum mecanismo para detectar mensagens vindas de um MTA *open relay*. Dentre várias opções comerciais e gratuitas, foi selecionado o projeto ORDB.org (http://www.ordb.org/), que disponibiliza uma *blacklist* de *open relays* através de consultas simples a um servidor de DNS (*Domain Name System*) contendo os endereços de MTAs mal-configurados.

4. RESULTADOS E DISCUSSÃO

Para a realização deste trabalho, foi selecionado um servidor com a distribuição Fedora Core 3, que possui cerca de 250 usuários ativos, além de várias listas de discussão internas e externas, alcançando cerca de 300 endereços de *e-mail*. Esse servidor executa a versão 2.1.5 do MTA *Postfix* junto com o *Postgrey*, versão 1.21 para habilitar o *greylisting*. Foram ativadas as opções do próprio *Postfix* para verificar a base de

1 Por envelope, entende-se a seqüência de comandos `MAIL From` e `RCPT To` do SMTP, que respectivamente identificam o remetente e o destinatário da mensagem [STEVENS, 1994].

dados do ORDB.org e bloquear *open relays*. Para verificar o desempenho dessa configuração, foram adotadas duas técnicas:

- avaliação dos registros (*logs*) do servidor de *e-mail* para verificar quantas mensagens foram bloqueadas e quantas foram efetivamente entregues, no período de 07/08/2005 a 08/09/2005;
- comparação do número de mensagens não solicitadas recebidas pelo primeiro autor deste trabalho[2], analisando o número de mensagens indesejadas antes e depois da adoção das técnicas aqui apresentadas.

A Tabela 1 mostra um resumo das informações levantadas nos registros do servidor. Os registros indicam que o total de 81.412 mensagens foram bloqueadas pelos dois mecanismos, sendo que 81.149 delas foram barradas através do *Postgrey*. Isso significa que mais de oitenta mil mensagens não foram retransmitidas pelo MTA do remetente. Se for assumido que todas as 126.423 mensagens constituem *spam*, houve uma redução de mais de 60% do volume original. Uma análise mais detalhada não é viável sem conhecimento do conteúdo das mensagens efetivamente entregues, ação que não é possível no contexto.

Tabela 1: Número de mensagens analisadas através dos registros.

Mensagens	**Número**
Bloqueadas pelo Postgrey	81.149
Bloqueadas pelo ORDB.org	263
Efetivamente entregues	45.011
Total	126.423

Complementando a análise, o gráfico da Figura 1 mostra o total de mensagens certificadas como *spam* pelo autor deste texto. Pelos dados do gráfico, pode-se notar que a média de *spam* recebido por um único usuário nos períodos de agosto–outubro/2004 e agosto–outubro/2005 permitem observar que a redução do número de mensagens não solicitadas é de 97,38%, bem acima do valor obtido pela análise dos *logs*. Convém destacar que os 2–3 *spams*/dia do período em análise também incluem mensagens vindas de contas de *e-mail* que não são protegidas pelo *Postgrey* e que são redirecionadas para o autor, o que indica que a redução efetiva pode ser maior que a apresentada.

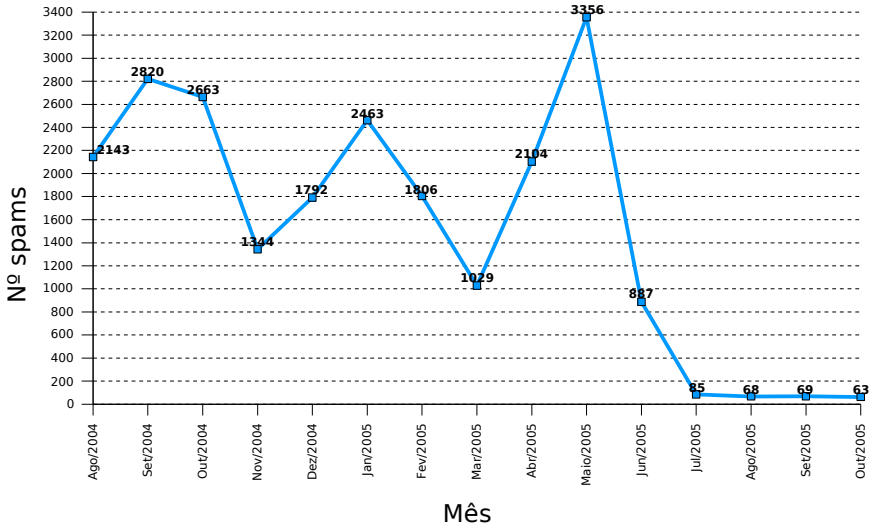

Figura 1: Número de mensagens indesejadas recebidas entre Agosto/2004--Agosto/2005.

2 Cabe aqui ressaltar que o autor salva todas as mensagens que manualmente identifica como *spam* desde agosto/2004, contabilizando mais de vinte mil mensagens que poderão ser usadas em treinamento de outros mecanismos *anti-spam*.

5. CONCLUSÕES

Controlar o *spam* não é uma tarefa simples. A cada técnica desenvolvida para combater o problema, novos meios de invalidá-la são encontrados. As técnicas usadas neste trabalho mostraram-se úteis na redução do número de mensagens recebidas pelos usuários, que manifestam, em informalmente, satisfação com o resultado. Não foram reportados casos de falso positivo ou negativo pelos usuários nem pelos registros feitos pelo *Postgrey*.

Um dos melhores resultados obtidos foi a postergação da aquisição de um novo servidor para abrigar as contas dos usuários e o serviço de *e-mail*. A máquina que atualmente provê o serviço não suportaria um mecanismo clássico baseado em análise de conteúdo com o volume original de mensagens.

Com um número reduzido de mensagens, observa-se que a adoção de uma técnica complementar de filtragem de *spam*, dentre as apresentadas na Seção 2, pode ser adotada com uma pequena melhoria no *hardware* em uso. Como trabalho futuro, propõe-se a avaliação dos impactos dessa filtragem local e a eficácia na redução de mensagens indesejadas, analisando o percentual de mensagens identificadas como *spam.*

REFERÊNCIAS

CARRERAS, X.; MÁRQUEZ, L. Boosting trees for anti-spam email filtering. In: *Proceedings of RANLP-01, 4th International Conference on Recent Advances in Natural Language Processing*. Tzigov Chark, BG: [s.n.], 2001.

CERT.BR. *Cartilha de Segurança para Internet, Part VI: Spam*. 2005. Disponível em: http://cartilha.cert.br.

CRISPIN, M. *RFC-1730 – Internet Message Access Protocol - Version 4*. 1994.

CUNNINGHAM, P. et al. A case-based approach to spam filtering that can track concept drift. In: *The CCBR'03 Workshop on Long-Lived CBR Systems, Trondheim, Norway*. [s.n.], 2003.

HARRIS, E. *The Next Step in the Spam Control War: Greylisting*. 2003. Disponível na internet. Último acesso: 09/09/2005. Disponível em: http://www.greylisting.org/articles/whitepaper.shtm.

MANCO, G. et al. Towards an adaptive mail classifier. In: *Italian Association for Artificial Intelligence Workshop Su Apprendimento Automatico: Metodi Ed Applicazioni*. [s.n.], 2002.

MYERS, J. G. *RFC-2554 — SMTP Service Extension for Authentication*. 1999.

MYERS, J. G.; ROSE, M. T. *RFC-1939 — Post office protocol version 3*. 1996.

NEMETH, E. et al. *UNIX System Administration Handbook*. 3. ed. New Jersey: Prentice-Hall, 2001.

ODA, T.; WHITE, T. Developing an immunity to spam. In: *Genetic and Evolutionary Computation Conference*. Chicago: Springer, 2003. v. 2723, p. 231–242.

ODA, T.; WHITE, T. Increasing the accuracy of a spam-detecting artificial immune system. In: *Proceedings of the Congress on Evolutionary Computation*. Canberra, Australia: [s.n.], 2003. v. 1, p. 390–396.

POSTEL, J. B. *Simple Mail Transfer Protocol, RFC 821*. Set. 1982. 68 p.

SAHAMI, M. et al. A bayesian approach to filtering junk E-mail. In: *Learning for Text Categorization: Papers from the 1998 Workshop*. Madison, Wisconsin: AAAI Technical Report WS-98-05, 1998.

STEVENS, W. R. *TCP/IP Illustrated: the protocols*. [S.l.: s.n.], 1994.

TRETYAKOV, K. Machine learning techniques in spam filtering. In: *Data Mining Problem-oriented Seminar, MTAT.03.177*. [s.n.], 2004. p. 60–79.

TURNER, D.; HAVEY, D. Controlling spam through lightweight currency. In: *Proceedings of the Hawaii International Conference on Computer Sciences, Jan 2004. 8*. [s.n.], 2004.