Uma solução de filtro anti-spam individualizado para ambientes de larga escala

Guilherme P. Pezzi, Francisco F. Fialho, Leandro F. Rey

Universidade Federal do Rio Grande do Sul
Centro de Processamento de Dados - CPD/UFRGS - Porto Alegre, RS
{guilherme, francisco, leandro}@cpd.ufrgs.br

Introdução

O Chasque Mail [MAR 2005] possui aproximadamente 15 mil usuários e, desde sua criação, utilizou-se um filtro anti-spam global e estático para todos usuários, através do DSPAM. Com o objetivo de melhorar a eficácia da marcação de spam, foi proposta a utilização de regras individuais para classificação de spam. O DSPAM oferece esse mecanismo, porém cada base individual de usuário pode ocupar entre 5 e 10 MB. Ao ativar as bases individuais para todos os usuários, observou-se um grande aumento do atraso na entrega das mensagens devido ao processamento do DSPAM e da manutenção do banco de dados. No entanto, não existe melhora significativa na eficácia do filtro apenas ativando as bases individuais, caso não haja um retorno do usuário. Por esse motivo, optou-se ativar bases individuais apenas para usuários que efetivamente treinarem o DSPAM. Este artigo apresenta uma solução desenvolvida para ativar a base individual de cada usuário no primeiro treinamento do DSPAM e as configurações utilizadas até se obter uma solução adequada.

Configurações do DSPAM utilizadas

A seguir, são descritas duas configurações utilizadas ao longo do tempo. Primeiro, uma configuração com uma base de regras dinâmica e global para todos usuários. Após, descreve-se a configuração para bases individuais e apresenta-se os tempos de atraso obtidos.

Base global dinâmica

Para permitir que os próprios usuários atualizem as regras de filtragem, o DSPAM é utilizado no modo `TOE` (Train on Error), com uma base única para todos usuários. Desse modo, quando o DSPAM errar a marcação de uma mensagem o usuário pode notificar o erro e melhorar as regras de filtragem. No entanto, cada usuário pode classificar diferentemente uma mesma mensagem, tornando difícil ter um conjunto de regras que seja adequado para todos usuários. Nessa configuração, a base de dados estabilizou-se com 5 GB e tempos de atraso menores do que 60s.

Bases individuais para todos usuários

Com o objetivo de individualizar as bases de dados para cada usuário e obter regras de filtragem específicas para cada usuário, utilizou-se o modo de treinamento TOE e a opção `Merged Groups` do DSPAM. Nesse modo, utiliza-se uma base pré-treinada em conjunto com as bases individuais. Isto significa que até que a base individual esteja consolidada, classifica-se as mensagens utili-

zando principalmente as regras da base pré-treinada, que é estática. Ao longo da fase de treinamento da base individual, as regras do usuário passam a fazer diferença na classificação das mensagens.

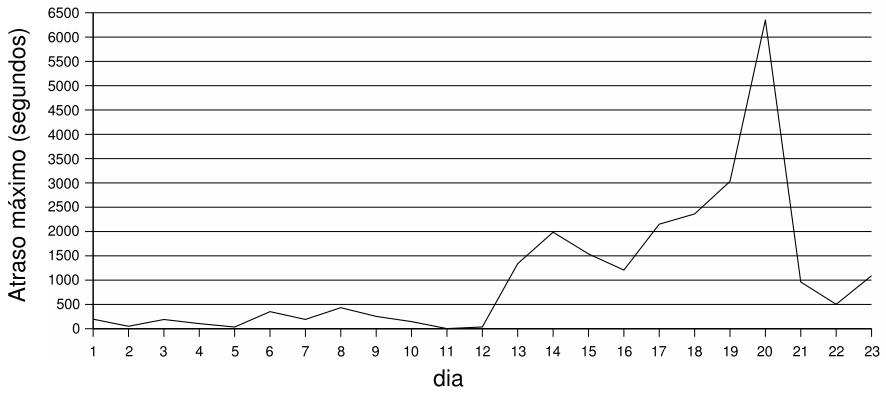

Figura 1: Tempos máximos de atraso pelo processamento do anti-spam.

A fase de treinamento inicial ocorre até que cada usuário receba 1000 mensagens. Nessa fase, o banco de dados MySQL aumenta rapidamente. Na Figura 1, apresenta-se os tempos máximos de atraso pelo processamento do DSPAM. Observou-se, em 3 semanas, que o tempo de atraso não se estabilizou e foi maior do que 1h, o que inviabiliza utilizar essa configuração para o filtro, pois tem um custo muito alto. Quando essa configuração foi descartada, o tamanho da base já era de 16GB.

Solução desenvolvida: bases individuais apenas para usuários que re-treinarem o filtro

Utilizou-se a configuração anterior para o DSPAM, porém foi desenvolvido um mecanismo que cria bases individuais apenas para usuários que notificarem erros do filtro. Foi adicionada uma tabela no MySQL que armazena quais usuários já executaram re-treinamento. Essa tabela é consultada na entrega da mensagem pela aplicação de entrega de e-mails (Maildrop) e atualizada pelo script de re-treinamento. Quando o usuário reporta pela primeira vez uma marcação errada do filtro, ele é adicionado à base MySQL. Quando o Maildrop recebe uma mensagem, consulta-se a tabela MySQL para definir os argumentos de execução do DSPAM: se usuário já retreinou utiliza-se base individual e, caso contrário, utiliza-se a base global estática. Após 2 meses, existem aproximadamente 380 bases individuais, o tempo de atraso em média é menor que 1s e o tamanho da base está em 2.4GB.

Para obter essa solução, foi desenvolvido um patch para o Maildrop e um script de retreinamento (perl), que atualiza a tabela MySQL de usuários com bases individuais. Maiores informações e os fontes dessa solução podem ser encontrados em [WIK 2006], na seção `Third Party Tools`.

Referências

[MAR 2005] MARCHI, A. et al. Chasque: o correio eletrônico da ufrgs migrando para o software livre. In: WORKSHOP SOBRE SOFTWARE LIVRE, 2005. **Anais...** [S.l.: s.n.], 2005.

[WIK 2006] WIKI dspam. Disponível em http://dspamwiki.expass.de/. Acessado em jan/07.