Relatório do Protótipo AntiSpam

Elvis Nobrega de Alcantara

25 de abril de 2015

1 – Introdução

Expressões regulares são uma facilidade em tratamento de textos em programas baseados em Perl. Podem ser usados padrões de substituição, pesquisa, até mesmo de tradução. Usarei tais métodos para implementação de um AntiSpam básico baseado em pesquisa de palavras chave e etc.

2 – Implementação do programa

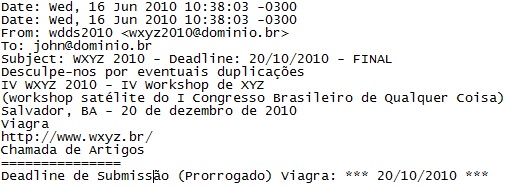
O programa foi programado em Perl, linguagem focada em textos, usando seus módulos strict e warnings, auxiliares no processo de debug.

O programa não encontra claras restrições, a não ser uma situação onde as palavras chaves são substituídas por equivalentes codificados como Viagra -> Vi@gra.

Basicamente o programa pega informações de um arquivo de configuração, a fim de conseguir colocar valor em limites do que é spam e o que não é, então usa essas configurações para verificar palavras chave no e-mail que podem defini-lo como um spam, e por fim mostra o resultado. De acordo como eu coloquei meu programa, se futuramente fosse possível guardar informação sobre e-mails ele poderia facilmente guardar suspeita em determinados e-mails, coisa que pode ser pesada, levando em conta que boa parte dos e-mails spam são enviados apenas uma vez.

3 – Casos de uso

Baseado na leitura de um e-mail de extensão .txt e um arquivo de configuração cujo nome pode ser escolhido pelo usuário, e tem como organização: comentários começam com #, em especial os comentários que definem o tipo de configuração que se seguirá. Sobre como o arquivo de texto do e-mail está organizado eu uso de base o e-mail de exemplo do professor:



4 – Conclusão

Muito instrutivo o serviço com expressões regulares, elas são muito úteis. Ideias novas podem ser facilmente adicionadas como métodos de outras chaves, ou outros e-mails, whitelist, etc. Minha intenção é fazer um código expansível.