

SpamOracle

Author: André Felipe Machado<clubes@techforce.com.br>

[SpamOracle](#) é um filtro anti-spam com algoritmo de Bayes para e-mails em português e outros idiomas.

É um [filtro anti-spam com algoritmo de Bayes](#) auto-aprendizado de formação de padrões a aceitar e a rejeitar. Você "mostra" para ele o que são SUAS mensagens boas e as que são spam e ele aprende a filtrar.

Quanto mais usa, mais eficaz fica, pois tem mais exemplos para análise.

Supera SpamAssassin em todos aspectos (velocidade, tamanho, praticidade, facilidade, você mesmo ensina o que é um spam novo que ele ainda não capturava) e não depende de alguém analisar o spam e criar uma nova regra. Criar regras de filtragem é uma caça em que os spammers estão sempre um passo à frente. Frustrante.

A partir de certo quantidade de emails exemplo, ele passa a ter ZERO falsos positivos. Falso positivo é um email "bom" erroneamente classificado como spam.

SpamOracle é o primeiro filtro com algoritmo Bayes que funciona em mensagens de idiomas diferentes do inglês (outubro 2002). Você pode configurar e compilar para aceitar apenas idiomas que você usa.

SpamOracle é o primeiro filtro com algoritmo Bayes capaz de analisar ANEXOS binários e textuais também.

SpamOracle pode ser usado com procmail (readme) para vários programas, ou SEM PROCMAIL no programa Kmail, filtrando contas IMAP e pop. É uma ótima opção para usuários desktop. ~~Site~~ [Site](#) crevemos um mini how to que foi aceito para inclusão na seção [tools](#) oficial do Kmail. Está em inglês, pois você PRECISA ler também o readme do SpamOracle.

Outros programas de email, que suportem a ação de filtrar mensagem através de programa externo, também podem ser usados. Lendo o mini how to provavelmente você conseguirá configurar estes programas também.

Para compilar o programa você precisará de um compilador Objective Caml. Você pode baixá-~~Existem pacotes prontos~~ para Linux e Windows. Se quiseres arriscar e usar o SpamOracle binário que compilei para pentium no cl8, configurado para aceitar inglês, português, espanhol e alemão, baixe [aqui](#) path ou chamar incluindo o path que escolheres. Ainda assim vais ter de ler o mini-how-to e o readme citados para poder configurar. Melhor mesmo é compilar e instalar, pois terá as man pages instaladas também e máximo desempenho em sua máquina.

Algumas distribuições linux, como Debian, já possuem pacote para o SpamOracle.

Se desejares acelerar o aprendizado de seu programa, podes já começar com uma grande amostra de spam correntes na internet brasileira.

Como nossos endereços estão na internet desde 96, e também publicados em sites, figuram em todos bancos de dados de spammers. Somos dos primeiros a sofrer cada nova onda de mensagens indesejadas. Se você tiver cópia atualizada dos spam recebidos, provavelmente a sua instalação do programa aprenderá mais eficazmente e apagará as mensagens antes de ter de gastar seu tempo a primeira vez. Naturalmente poderás adicionar e remover suas próprias mensagens para aprendizado, se necessário. Afinal, você pode ter um conceito do que classificar como spam diferente do nosso.

Para receber cópias atualizadas regularmente dos spam que recebemos, por R\$ 10 mensais, fale conosco.