"Democratizando" a Filtragem e Bloqueio de Conteúdos Web

Filipe Pires[1], Alexandre Fonte[1], Vasco Soares[1]

[1] Escola Superior de Tecnologia de Castelo Branco
Instituto Politécnico de Castelo Branco
Av. do Empresário, 6000-767 Castelo Branco, Portugal.

{fpires,adf,vasco_g_soares}@est.ipcb.pt

Resumo

A filtragem e bloqueio de conteúdos Web é um assunto polémico e controverso. Para isto contribui o facto de que a maioria dos sistemas que efectuam esta actividade se encontram maioritariamente implementados em países com regimes políticos opressivos, sob a forma de mecanismos legais e tecnológicos de censura. Tal mediação vai contra os princípios gerais da Internet, uma rede global de partilha de informação pública, revogando os direitos dos utilizadores face à utilização de tais sistemas. Estes sistemas encontram-se actualmente numa fase de proliferação e existem certas áreas onde a sua aplicação se poderá tornar benéfica. Um exemplo destas áreas é a filtragem e bloqueio de conteúdo pedófilo. Neste artigo apresenta-se a arquitectura de um sistema de filtragem e bloqueio de conteúdos Web, denominado Sisbloque, projectado para ser implementado sobretudo em ISPs (Internet Service Providers), grandes instituições ou companhias, que para além de possuir um conjunto melhorado de mecanismos de filtragem de conteúdos, introduz um conceito inovador ao nível de transparência suportado pelo seu mecanismo de manipulação de erros.

1 Introdução

A Internet é cada vez mais um ambiente inseguro quer para partilha de informação quer em qualquer outra actividade, tal deve-se ao facto da criminalidade se começar a integrar de forma consistente nas novas tecnologias. Apesar das controvérsias que a filtragem e o bloqueio de acesso a conteúdos Web suscita, este método apresenta-se como uma solução eficaz frente a problemas como a publicação on-line de conteúdos pedófilos. Actualmente a maioria dos sistemas de filtragem e bloqueio de conteúdo Web, ou são soluções proprietárias ou produtos comerciais, sendo a maioria dos detalhes de implementação destes sistemas desconhecidos pela comunidade científica.

Esta lacuna motivou o desenvolvimento e concepção de um sistema aberto de filtragem e bloqueio de conteúdos Web, designado por Sisbloque. Este sistema está concebido para potencial uso em ISPs, grandes companhias ou instituições que necessitem deste tipo de serviço e propõe um mecanismo de filtragem de conteúdos mais conciso e fiável, proveniente do melhoramento de métodos existentes como a filtragem baseada na origem, filtragem de conteúdos e imagens.

O restante conteúdo deste artigo encontra-se organizado da seguinte forma. A secção 2 descreve a arquitectura do sistema Sisbloque. A secção 3 apresenta os resultados referentes a alguns testes de desempenho efectuados ao protótipo do sistema Sisbloque. Finalmente, a secção 4 conclui este artigo.

2 Visão Geral da Arquitectura Sisbloque

O sistema Sisbloque é composto por três módulos distintos que interagem entre si: o módulo de filtragem de conteúdos Web, o módulo de serviços e o módulo de manipulação de erros (ver figura 1) [1].

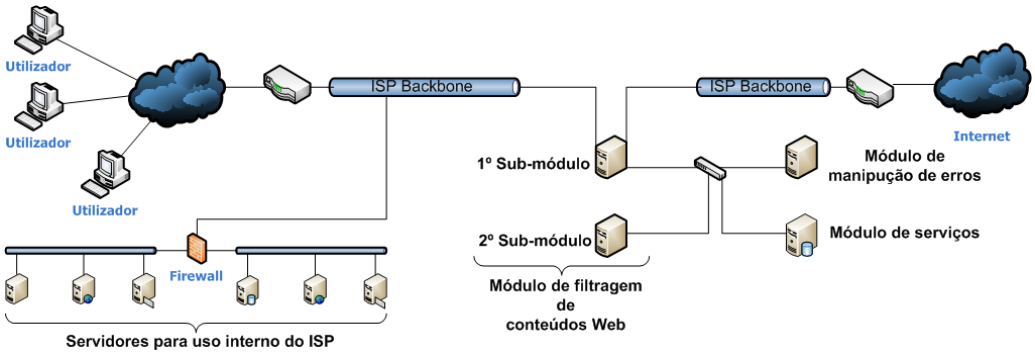

Figura 1: Visão geral da Arquitectura Sisbloque.

O módulo de filtragem de conteúdos Web é composto por dois sub-módulos: o primeiro consiste num método de filtragem baseado na origem do tráfego; o segundo é composto por um filtro de conteúdos mais específico. Esta abordagem permite tirar maior partido das vantagens de cada método, sem comprometer o sistema com as desvantagens de cada um.

O módulo de serviços é composto por todos os servidores necessários ao funcionamento do sistema, como por exemplo o suporte ás bases de dados que contêm as diversas listas de endereços, sejam estas a lista de inclusão, a lista de exclusão e lista de imunidade.

O módulo de manipulação de erros é responsável por elaborar uma mensagem de erro real, para que esta seja retornada a um utilizador sempre que este tenta aceder a um web site malicioso.

Finalmente, é importante notar que o desenho desta arquitectura procurou considerar três requisitos fundamentais, os quais devem ser observados pelo sistema Sisbloque; a saber um baixo custo de implementação e manutenção, uma elevada fiabilidade e precisão na avaliação de conteúdos Web, bem como um elevado grau de transparência.

Nas subsecções 2.1.1 e 2.1.2 são apresentados em detalhe o mecanismo de filtragem de conteúdos e o mecanismo de manipulação de erros. Na secção 2.2 é discutido uma característica chave do sistema Sisbloque, que é a sua potencial modularidade.

2.1 Tratamento de Conteúdos

2.1.1 Mecanismo de Filtragem de Conteúdos

O filtro baseado na origem do Sisbloque é o primeiro filtro do mecanismo de filtragem de conteúdos [2]. No sistema Sisbloque este filtro é constituído por duas técnicas de filtragem, uma orientada aos endereços URL contidos no protocolo HTTP e outra baseada nos endereços IP dos respectivos pacotes de dados. Estes endereços são posteriormente avaliados pelo filtro de inclusão e exclusão.

O filtro de inclusão funciona através de uma lista de acesso composta por endereços URL/IP autorizados [2-3]. Os endereços URL/IP contidos nesta lista de acesso contêm apenas informação segura, relativa a instituições de ensino, bancos, serviços governamentais, entre outros, sendo criada e mantida pelo administrador do sistema.

O filtro de exclusão funciona através de uma lista de bloqueio composta por endereços URL/IP banidos [2]. A lista de bloqueio do sistema é actualizada sempre que o filtro de conteúdos detecta um novo servidor Web malicioso, ou quando um servidor Web de conteúdo malicioso muda de endereço IP. Noutros sistemas, as listas de bloqueio são actualizadas por entidades externas, como IWF (Internet Watch Foundation) ou ECPAT (End Child Prostitution, Child Pornography and Trafficking of Children for Sexual Purpose) [4-5]. Contudo este procedimento pode induzir erros nas respectivas listas, como a introdução de servidores

Web não maliciosos, devido a falsas denúncias. De modo a garantir uma menor susceptibilidade da lista de bloqueio a ataques DNS [2], deve configurar-se no sistema Sisbloque um conjunto fiável de servidores DNS, para que este possa comparar os dados provenientes de vários servidores DNS em simultâneo.

Existe ainda o filtro de imunidade, o qual possui uma lista de endereços URL/IP referentes a web sites que não deverão ser incluídos nem na lista de acesso nem na lista de bloqueio. Tal deve-se ao facto de determinados web sites que apesar de poderem retornar conteúdo malicioso, não são directamente responsáveis por ele. Um exemplo deste caso são os motores de busca e também web sites de alojamento, entre outros. Este método garante que todo o conteúdo proveniente destes web sites será sempre convenientemente filtrado.

O filtro de conteúdos Web arquitectado para o sistema Sisbloque, é constituído por uma técnica de filtragem mais específica do que a habitual filtragem por palavras [1]. Sempre que um web site não se encontre incluído na lista de acesso, na lista de bloqueio e na lista de imunidade, o seu conteúdo é comparado as duas listas distintas, uma composta por palavras caracteristicas de web sites maliciosos e uma composta por palavras não maliciosas, onde a cada palavra maliciosa é atribuído um valor positivo e a cada palavra não maliciosa um valor negativo. Posteriormente é comparado a cotação geral do web site com o valor limite predefinido pelo administrador de sistema. Caso a cotação seja mais elevada que o limite definido então o web site em causa é considerado malicioso, sendo de imediato adicionado à lista de bloqueio do sistema.

2.1.2 Mecanismo de Manipulação de Erros

Em geral de entre os sistemas de filtragem e bloqueio de conteúdos Web que prepondera m nesta área, quando é bloqueado o acesso de um utilizador a determinado conteúdo, é apresentado ao utilizador um aviso ou em alguns casos este é simplesmente deixado sem resposta. Este tipo de respostas torna perceptível, inclusivamente a utilizadores menos experientes, que algo está a bloquear o seu acesso aos conteúdos.

O grau de transparência neste tipo de sistemas é um factor de extrema importância, pois caso a metodologia de como o sistema reage quando permite ou bloqueia o acesso a conteúdos Web seja perceptível, então o sistema em causa encontrar-se-á vulnerável a possíveis ataques Oracle [6]. Segundo Lowe [6], um ataque Oracle consiste em fazer com que um sistema responda rigorosamente a qualquer número de perguntas que lhe são efectuadas, sem ter a noção das possíveis consequências. Se um sistema deste âmbito for vulnerável a este tipo de ataque, é então possível efectuar um scan de um intervalo de endereços IP, e de acordo com o tipo de resposta obtido é então construída uma lista semelhante, se não exactamente igual, à lista de bloqueio presente no respectivo sistema de filtragem. Este tipo de ataque foi efectuado e documentado sobre o sistema Cleanfeed [7].

No sistema Sisbloque sempre que um utilizador efectua um pedido de acesso a conteúdos maliciosos, este é redireccionado para o mecanismo de manipulação de erros. O mecanismo de manipulação de erros do sistema Sisbloque gera aleatoriamente erros da gama 5xx do protocolo HTTP. Posteriormente cada vez que um utilizador tenta aceder a conteúdo malicioso, é apresentado um erro gerado pelo mecanismo de manipulação de erros, iludindo o utilizador perante uma falha referente ao servidor de conteúdos a que este tentou aceder, garantindo portanto um nível mais elevado de transparência ao sistema. Através deste mecanismo torna-se imperceptível ao utilizador se o erro em causa é proveniente do servidor de conteúdos ou do sistema Sisbloque, tal garante uma maior robustez por parte do sistema a ataques deste tipo.

2.2 Modularidade do Sistema

No que diz respeito à capacidade de modularidade do sistema Sisbloque, este segue uma abordagem de implementação de um sistema distribuído. Como já foi acima referido, o

sistema Sisbloque é composto por três módulos distintos sendo estes: módulo de filtragem de conteúdos Web, módulo de serviços e módulo de manipulação de erros (ver figura 1). Destes três módulos é denotar ainda que o módulo de filtragem de conteúdos Web se divide em dois sub-módulos. Assim o sistema pode ser distribuído em vários suportes físicos, dedicando o poder de processamento e os recursos disponíveis de cada suporte físico exclusivamente a um único módulo ou sub-módulo. Através desta implementação obtém-se maior desempenho por parte de cada módulo, o que aumenta significativamente a capacidade de resposta geral do sistema em si. Dependendo dos recursos disponíveis por parte da empresa ou instituição, fica ao critério do administrador do sistema gerir a distribuição dos diferentes módulos, podendo em último recurso integrar todos os módulos em apenas um suporte físico se assim for necessário.

3 Protótipo do Sistema Sisbloque

O protótipo do sistema Sisbloque é suportado por um conjunto de componentes de software abertos, amplamente disponíveis e de uso livre, como tal a sua implementação baseia-se na integração destes componentes e de um conjunto de extensões e melhoramentos introduzidos à medida. Este conjunto tem como ambiente base de execução o sistema operativo Linux, mais especificamente a distribuição Fedora Core 8, e a respectiva kernel 2.6.25.9 a qual fornece suporte a uma framework de filtragem de pacotes designada de Netfilter [8]. A gestão de acessos a endereços URL/IP é implementada com base na integração do Web proxy Squid, juntamente com o plugin SquidGuard [9-10]. Um servidor HTTP de tecnologia Apache compõe o sistema com o propósito do Sisbloque poder usar e jogar com determinados parâmetros do protocolo HTTP durante a sua actividade de bloqueio do tráfego HTTP [11]. Para suporte de informação é usado o sistema de gestão de bases de dados relacionais, MySQL [12].

Na secção 3.1 são discutidos os detalhes do actual protótipo do sistema. Na secção 3.2 são apresentados os resultados de uma avaliação de desempenho ao protótipo.

3.1 Detalhes de Implementação

O módulo de filtragem de conteúdos Web, sendo constituído por dois sub-módulos, é o modulo mais complexo do sistema Sisbloque. O primeiro sub-módulo é introduzido num ponto específico da topologia de rede, onde passa obrigatoriamente o fluxo de dados a ser filtrado, sendo a prévia ligação existente restabelecida através de uma bridge de rede controlada pela framework Netfilter. O controlo de acessos a conteúdos Web, é também efectuado neste sub-módulo, através do Web proxy Squid, onde os pedidos a endereços IP/URL contidos na lista de bloqueios são redireccionados para o módulo de manipulação de erros. Durante este processo, é ainda enviado ao segundo sub-módulo qualquer endereço que não se encontre contido na lista de bloqueio, de acesso ou de imunidade. Cabe ao segundo sub-módulo avaliar os endereços recebidos, através do filtro de conteúdos Web. Caso seja determinado que um endereço possui conteúdo considerado malicioso, este é adicionado à lista de bloqueio do sistema que se encontra no módulo de serviços, sendo posteriormente enviado um pedido de actualização ao primeiro sub-módulo para que este actualize a sua lista de bloqueio.

No módulo de manipulação de erros reside o mecanismo de manipulação de erros. Este módulo tem como objectivo base a geração de códigos de erro aleatórios da gama 5xx do protocolo HTTP. Para tal é utilizado neste módulo o servidor HTTP Apache, o qual é periodicamente forçado a causar de forma aleatória os erros pretendidos. Sempre que um erro é gerado existe um tempo de duração que lhe é associado, isto permite que vários utilizadores não se deparem com erros diferentes no mesmo intervalo de tempo quando solicitarem conteúdo considerado malicioso.

O módulo de serviços serve de suporte ao sistema de gestão de bases de dados relacionais MySQL. É neste módulo que as diferentes listas do sistema são armazenadas, de forma a manter a sua coerência de dados e também a facilitar a sua actualização, sempre que um módulo de filtragem de conteúdos Web efectua uma actualização ou inicializa o seu processo.

De notar que de entre estes módulos, apenas o módulo de filtragem de conteúdos Web interage directamente com a rede. Os restantes módulos encontram-se isolados da rede principal, sendo apenas acessíveis pelo mesmo módulo de filtragem de conteúdos Web.

3.2 Avaliação de Desempenho

O protótipo do sistema Sisbloque encontra-se em fase de desenvolvimento, durante a qual têm sido efectuados testes ao sistema. Nesta secção discute-se os resultados obtidos relativamente à avaliação de desempenho ao qual o sistema foi submetido.

A avaliação de desempenho foi efectuada ao primeiro sub-módulo do módulo de filtragem de conteúdos Web. O seu objectivo foi consignar os recursos usados pelo respectivo módulo (ver figura 2) bem como as latências induzidas nos utilizadores (ver figura 3), quando submetido o sistema a uma sobrecarga de informação a filtrar. Este sub-módulo é composto por um processador Pentium 4 3.0GHz, por 1Gb de memória DDR 400MHz TwinMOS, por uma motherboard Gigabyte GA-8I915G-MF sendo o seu chipset Intel 915G Express.

A avaliação de recursos focou-se na percentagem de processamento usado pela unidade central de processamento, dividindo-se esta na percentagem de utilização de CPU por parte dos processos e na percentagem de utilização de CPU por parte do sistema operativo. A avaliação de latência determinou a latência provocada nos utilizadores da rede, à medida que o número de estações, que originaram a sobrecarga, foi incrementado.

No que diz respeito à sobrecarga, foi elaborado um ficheiro contendo aleatoriamente cinco mil endereços de servidores de conteúdos Web considerados maliciosos e cinco mil endereços de servidores de conteúdos Web considerados não maliciosos, denotar que nenhum destes endereços foi repetido, posteriormente este ficheiro foi dividido em vinte ficheiros os quais foram distribuídos por vinte estações. A sobrecarga teve um período de sessenta minutos no qual as vinte estações acederam ciclicamente, de um em um segundo, e em simultâneo ao conteúdo Web referente aos endereços incluídos nos respectivos ficheiros, originando desta forma tráfego HTTP a ser filtrado. Face a esta sobrecarga o sub-módulo em causa teve um aumento na sua percentagem de processamento total usado, permanecendo noventa e quatro porcento da sua capacidade de processamento disponível, foi também verificado um aumento relativamente à percentagem de memória usada contudo não foi de modo algum significativo. Durante este período foi ainda verificado que o tempo de acesso por parte dos utilizadores desta rede ao seu respectivo conteúdo sofreu alguma latência, a qual permaneceu pela média de sessenta e cinco milissegundos.

O tempo e os recursos utilizados pelo sub-módulo quando se efectua actualizações na lista de regras do Netfilter, varia consoante a localização onde o sistema é implementado na topologia da respectiva rede. A tabela de regras do Netfilter é composta principalmente pelas próprias regras que salvaguardam o sistema bem como pelas regras que reencaminham o tráfego relativo ao protocolo HTTP para o porto do Squid, adicionalmente e dependendo da localização do sistema na topologia de rede são adicionadas excepções correspondentes aos servidores internos da rede. Este número de excepções pode ser bastante reduzido ou praticamente inexistente caso o sistema seja integrado entre um servidor proxy e o respectivo gateway da rede em causa. Contudo, se esta lista possuir grande dimensão, então a sua actualização irá demorar algum tempo o que levará à indução de latências nos utilizadores, um método de evitar esta situação será a redireccionar o tráfego a ser filtrado para um segundo sub-módulo de filtragem isto enquanto o primeiro actualiza as novas definições, tornando a redireccionar o tráfego para o primeiro sub-módulo após a conclusão da actualização deste.

No respectivo protótipo do sistema a actualização quer da lista de regras do Netfilter

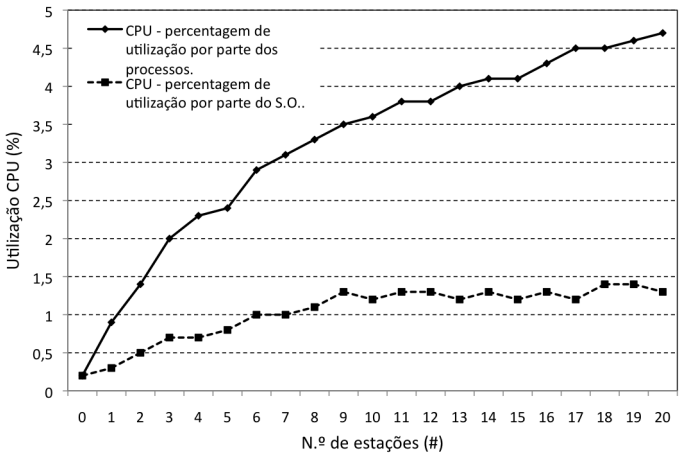

Figura 2: Percentagem de utilização de CPU.

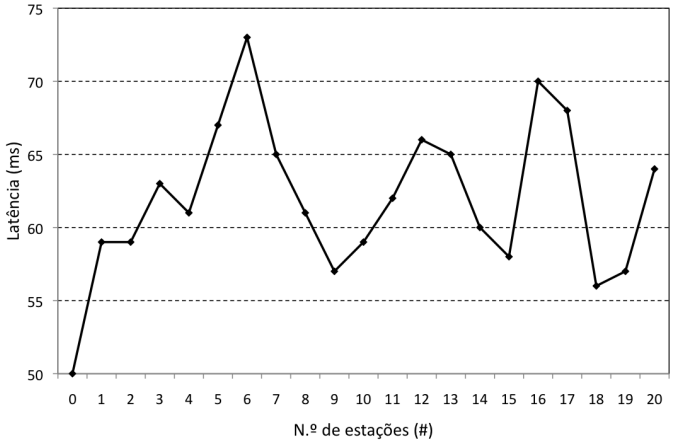

Figura 3: Latências induzidas nos utilizadores.

quer da lista de endereços do SquidGuard, embora requeira algum tempo de processamento tal não induz latência no acesso dos utilizadores.

4 Conclusão

Presentemente a criminalidade tira cada vez mais partido das novas tecnologias. Como tal, apesar da actual controvérsia em torno de sistemas de filtragem e bloqueio de conteúdos Web, estes revelam-se eficazes quando aplicados a determinados conteúdos como é o caso de conteúdos pedófilos ou qualquer conteúdo implícito a abuso de menores.

Neste artigo apresentámos a arquitectura integral de um sistema de filtragem e bloqueio desenvolvido para ser implementado sobretudo em ISPs. O Sisbloque tem como principais objectivos garantir um baixo custo de implementação e manutenção, associado a uma fiabilidade e a um sistema de filtragem e bloqueio de conteúdos Web altamente preciso. A sua arquitectura juntamente com os seus mecanismos de filtragem melhorados e o seu inovador

mecanismo de manipulação de erros, introduzem um novo conceito de transparência, o que apresenta melhorias significativas quando comparado a outros sistemas concorrentes.

Face à avaliação do protótipo do sistema concluímos que apesar do ambiente desta avaliação ser relativamente reduzido, o protótipo do sistema demonstra já uma excelente capacidade de resposta, contudo no melhor do nosso conhecimento não existe informação científica disponível relativa a testes de desempenho efectuados a sistemas concorrentes, o que impossibilita a sua comparação com o protótipo do sistema Sisbloque. O desenvolvimento deste projecto irá oferecer à comunidade científica informação única naquilo que é filtragem e bloqueio de conteúdos Web.

Referências

1. Pires, F., Fonte, A., Soares, V., "A Filtragem e Bloqueio de Conteúdos Web Segundo o Projecto Sisbloque". 3ª Conferência Ibérica de Sistemas e Tecnologias de Informação. pp. 1-6, 2008.

2. Greenfield, P., Rickwood, P., Cuong Tran, H., "Effectiveness of Internet Filtering software products", CSIRO Mathematical and Information Sciences, pp. 6-12, 2001.

3. Carlzon, M., Hagsand, O., Widell, F., Danielsson, B., "Blocking Web Contents using BGP", Royal Institute of Technology. Stockholm, Sweden, pp. 2-5, 2005.

4. Internet watch foundation, "The Internet Watch foundation", accessed at 5 October 2008, <http://www.iwf.org.uk/>.

5. Ecpat, "Ecpat sverige", accessed at 5 October 2008, <http://www.ecpat.se/>.

6. Lowe, G., " An Attack on the Needham-Schroeder Public-Key Authentication Protocol", Information Processing Letters, 56(3), p.131-133, 1995.

7. Clayton, R., Failures in a Hybrid Content Blocking System. Workshop on Privacy Enhancing Technologies. Dubrovnik, Croatia, p. 12, 2005.

8. Netfilter, "Netfilter/Iptables Project Homepage ", accessed at 5 October 2008, <http://www.netfilter.org/>.

9. Squid, "Squid-cache.org - Optimizing Web Delivering", accessed at 5 October 2008, <http://www.squid-cache.org/>.

10. Squidguard, "Squidguard", accessed at 5 October 2008, <http://www.squidguard.org/>.

11. Apache, "The Apache Software Foundation", accessed at 5 October 2008, <http://www.apache.org/>.

12. MySQL, "MySQL - The world's most popular open source database", accessed at 5 October 2008, <http://www.mysql.com/>.