Desenvolvimento de ferramenta para processar *logs* de *firewall*

Jonathan O. Preus, Sadan E. M. Figueira, Renato Preigschadt

[1]Colégio Técnico Industrial – Universidade Federal de Santa Maria (UFSM)
Av. Roraima, 1000, Cidade Universitária – 97.105-900 – Santa Maria – RS – Brazil

{jonathan.preus,sadan.figueira, renato}@redes.ufsm.br

Abstract. *The processing of log files becomes complex task due to the nature of the information provided, which, to maintain the system that generated efficient, must contain information on a technical level. However, the processing of these information can help in the hour of decision making by the network administrator. In this article we seek to develop a system capable of making a deep analysis of log files.*

Resumo. *O processamento de arquivos log se torna complexo devido à natureza das informações providas, as quais, para manter o sistema que a gerou eficiente, devem conter informações a nível técnico. Entretanto o processamento dessas informações pode ajudar na hora da tomada de decisões por parte do administrador da rede. Nesse artigo buscamos elaborar um sistema capaz de fazer uma profunda analise de arquivos de log.*

1. Introdução

A análise dos *logs* em busca de pacotes bloqueados, não permite uma análise rápida e eficiente em busca de informações essenciais como, por exemplo, entidade responsável pelo IP, serviços que são executados no *host*, (Skype, Facebook, WhatsApp ,etc).

Diante da complexidade de analisar e obter esse tipo de informações de *logs* de *firewall*, este trabalho propõe uma ferramenta com o propósito de gerar relatórios sobre as requisições negadas pelo *firewall*. Estes relatórios produzem informações relevantes e necessárias para a tomada de decisões para o administrator da rede, como endereços IP e portas, entidades responsáveis pelo endereço IP, existência de notificações de segurança em relação ao *host*, dentre outros.

Este trabalho está organizado da seguinte forma: na Seção 2 são apresentados conceitos básicos sobre firewalls e logs; na Seção 3 é apresentada a ferramenta desenvolvida para a análise de logs; a Seção 4 são apresentados e discutidos experimentos realizados com a ferramenta, e por fim na Seção 5 são apresentadas as conclusões deste trabalho.

2. Referencial Teórico

2.1. Firewall

Firewall é um componente ou conjunto de componentes que restringe o acesso entre uma rede protegida e a Internet, ou entre outros conjuntos de redes (Chapman et al 1995). O dispositivo atuante como *firewall* é responsável por analisar o tráfego de entrada e saída em uma rede e é baseado em uma série de políticas de controle de acesso, permitindo ou não um determinado tráfego. Para fins de registro histórico do

acionamento dessas políticas, é possível armazenar o resultado da execução das mesmas em um arquivo de *log*.

2.2. IBM X-Force Exchange

Em meados de 2015 a IBM lançou a IBM X-Force Exchange, esta plataforma armazena dados sobre endereços IPs, como domínios, responsáveis e se o endereço está ligado a incidentes de segurança entre outras informações. Todas essas informações são submetidas pelos próprios usuários, gerando assim uma base de dados constantemente atualizada sobre possíveis ameaças. Além dos recursos citados ainda possui uma API online a qual pode ser facilmente implementada. Para utilizar os recursos da API REST e integração da mesma com a aplicação desenvolvida, é necessário possuir cadastro no site da IBM X-Force.

Entre os vários tipos de informações que podem ser consultadas, as que mais se mostraram relevantes para o objetivo da aplicação desenvolvida foram:

- *whois*, que retorna as informações referentes aos endereços IPs tais como, data de criação dos registros, nome e organização a qual foi cadastrado, e-mail de contato, localização entre outros;
- *Ipr*, ou *ip repor*, que recupera informações mais detalhadas, entre elas o *score* do IP, uma métrica criada pelos usuários para mostra o grau de risco provenientes dos endereços IPs, sendo 1 para um IP de baixo risco e 10 para um endereço IP com baixa reputação que pode oferecer riscos;
- *ipr/malware*, serve para recuperar na base se existe alguma *malware* associado a esse endereço IP.

2.3. Trabalhos relacionados

Fábio Elias Locatelli, 2004, Propõem uma ferramenta para análise de logs em busca de identificar incidentes de segurança, no trabalho é descrito que cada evento gerado pelo firewall (Iptables), é comparado com um caso previamente estabelecido, presente em uma base de dados.

3. Desenvolvimento

3.1 Estruturas de mensagens de *log*

A estrutura das mensagens que registram os *logs* de ações do firewall apresenta um conjunto de informações sobre o resultado e os elementos envolvidos na execução das políticas de segurança. Os campos relevantes para a aplicação são: IN e OUT, que são referentes às interfaces de entrada e saída nas quais um determinado pacote pode trafegar. Os campos SRC e DST contêm os endereços IP’s de origem e destino dos pacotes, respectivamente. O campo PROTO indica o protocolo utilizado no pacote e os campos STP e DPT que se referem às portas de origem e destino da conexão.

3.2 Processamentos das mensagens de *log*

Com o objetivo de localizar somente as mensagens referentes às ações em que o *firewall* realiza o bloqueio de alguma requisição de conexão, foi construído um algoritmo para realizar a leitura de todo o arquivo de log em busca das linhas referentes. O fluxograma exibido na Figura 15, demonstra o funcionamento do algoritmo.

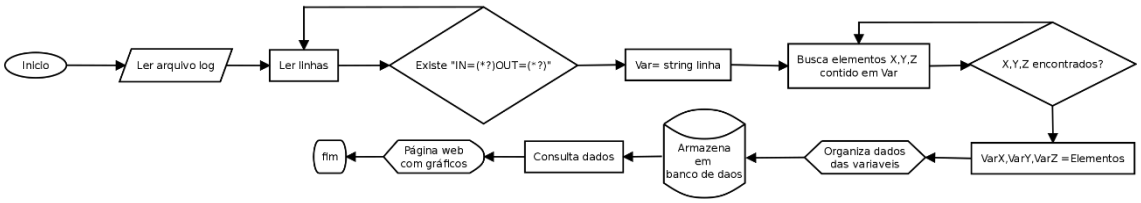

Figura 15 - Fluxograma para processamento de mensagem do log

O arquivo de *log* é apontado para o algoritmo pelo usuário, então usando expressões regulares o algoritmo percorrerá cada linha do *log* até encontrar as linhas que estejam em conformidade com a expressão regular. Ao encontrar essa linha, a mesma é atribuída a uma variável temporária para que possa ser tratada e selecionadas informações mais relevantes sobre a ação do firewall, como endereços de origem e destino do pacote, portas de comunicação, MAC e interfaces envolvidas na conexão. Cada elemento também é atribuído a uma variável para serem tratados e posteriormente inseridos em uma base de dados. Um exemplo das expressões regulares utilizadas para filtrar os dados de cada mensagem de log, é exemplificada na Figura 2. Esse trecho é utilizado para encontrar os endereços de ip de destino do pacote que tentou ser transmitido.

```
ObjSRC = re.search (r'SRC=(.*?)DST=(.*?)', line, re.M | re.I)
DST = ObjDST.group(2)
```

Figura 2 – Exemplo de expressão regular utilizada no sistema

3.4. Consumo da API de RESTFULL

A partir dos endereções IPS contidos na base de dados, é possível selecionar um ou múltiplos endereços IP para serem consultados pela API X-Force. Na aplicação proposta o uso da API se dá de acordo com o fluxograma representado na Figura 2.

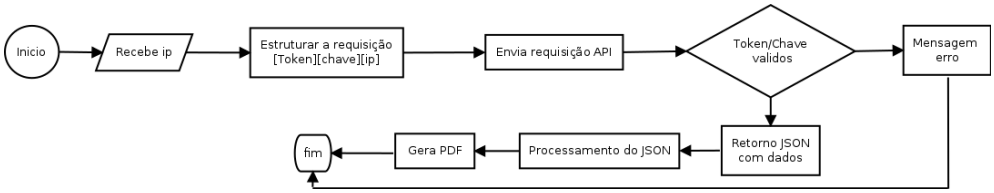

Figura 16 - Fluxograma de integração entre aplicação e API

Os dados de endereço IP, uma chave e um *token*, são enviados em uma requisição para API, aonde é verificado se a chave e o *token* são validos. Caso sejam, é retornado um JSON com os dados relacionados ao IP consultado. Esse JSON é processado pela aplicação desenvolvida e gera um arquivo de relatório.

4. Experimentos

Para os experimentos realizados foi utilizado os registros de *log* de um firewall, Netfilter/Iptables, instalado sobre em uma distribuição Linux Debian 8, que por sua vez está operando sobre um servidor, Intel(R) Xeon(R) CPU X3320 @ 2.50GHz, com 8GB de memória RAM e 92Gb de armazenamento. O *log* é referente ao período do dia 3 de maio, iniciando as 06:25:09 horas da manhã, até as 06:25:07 horas da manhã do dia 27 de abril de 2016. Após a primeira etapa de pré-processamento desse arquivo, foi gerado um conjunto de dados de 383007 (trezentos e oitenta e três mil e setes) linhas de registros. Todos esses registros se referem a tentativas de conexões bloqueadas pelo firewall.

Em um segundo momento, foi realizado um segundo processamento sobre esse novo conjunto de dados, para contabilizar os endereços IP distintos. Foram encontrados 4176 (quatro mil cento e setenta e seis) endereços distintos, de posse desses foi executada a aplicação para consulta, submetendo para a API X-Force os endereços e gerado um breve relatório dos mesmos. Desses endereços, a maioria não possuía registro ou alertas de periculosidade ou informações sobre os serviços servidos por eles. Porem também foram identificados endereços bloqueados, referentes a serviços como o AVAST Cloud, Skype, WhatsApp entre outros, somando um total de 526 (quinhentos e vinte e seis) ações de bloqueios. A figura 3, apresenta um relatório gerado sobre um ip.

```
Relatório sobre ip: 91.190.218.69, Data de criação: informação não disponível
Informações sobre o contratante: não disponível E-mail: informação não disponível
Tipo de contrato: registrant
Nome do contratante: Skype Du Organização responsável: informação não disponível
Pais do responsável: Ireland Incidentes: Informação não disponível
```

Figura 4 – Exemplo de relatório gerado

5. Conclusões

A ferramenta desenvolvida tornou a análise detalhada dos arquivos de log mais fácil. Permite realizar a análise sobre o que é cada ocorrência, quem são os responsáveis pelo endereço IP envolvido e se existe algum relato de incidentes de segurança referente a esse endereço.

Dessa maneira foi possível identificar que estavam sendo bloqueados equivocadamente pelo firewall endereços referentes aos serviços AVAST Cloud, Skype, WhatsApp, serviços estes importantes para os usuários.

As próximas atualizações da aplicação serão voltadas a melhorar sua usabilidade, tanto como o desenvolvimento de uma interface web que apresentará estatísticas dos serviços bloqueados e ameaças que requerem maior atenção, e aumentaremos sua confiabilidade fazendo consultas a um número maior de bancos de dados de IPs.

Referências

D. Brent Chapman, Elizabeth D. Zwicky, Simon Cooper. (1984), Building Internet Firewalls, 2nd edition.

API IBM X- Force, http://www.ibm.com/security/xforce/, acessado em 20/08/2016.

Identificação de Cenários de Intrusão pela Classificação, Caracterização e Análise de eventos gerados por Firewalls, http://ce-resd.facom.ufms.br/sbrc/2004/069.pdf, acessado em 22/08/2016.