Universidade Federal do Ceará

Centro de Tecnologia

Departamento de Engenharia de Teleinformática

Curso de Engenharia de Computação

BRUNO RICCELLI DOS SANTOS SILVA

IMPLEMENTAÇÃO E ANÁLISE DE UM FRAMEWORK DE DETECÇÃO DE ATAQUES DISTRIBUÍDOS DE NEGAÇÃO DE SERVIÇO

Fortaleza, Ceará

2017

BRUNO RICCELLI DOS SANTOS SILVA

IMPLEMENTAÇÃO E ANÁLISE DE UM FRAMEWORK DE DETECÇÃO DE ATAQUES DISTRIBUÍDOS DE NEGAÇÃO DE SERVIÇO

Monografia apresentada ao Curso de Engenharia de Computação da Universidade Federal do Ceará, como requisito parcial para obtenção do Título de Bacharel em Engenharia de Computação.

Orientador: Prof. Msc. Ricardo Jardel Nunes da Silveira

Co-Orientador: Prof. Msc. Marcelo Araújo Lima

Fortaleza, Ceará
2017

BRUNO RICCELLI DOS SANTOS SILVA

IMPLEMENTAÇÃO E ANÁLISE DE UM FRAMEWORK DE DETECÇÃO DE ATAQUES DISTRIBUÍDOS DE NEGAÇÃO DE SERVIÇO

Monografia apresentada ao Curso de Engenharia de Computação da Universidade Federal do Ceará, como requisito parcial para obtenção do Título de Bacharel em Engenharia de Computação.

Aprovada em: ___/___/_____

BANCA EXAMINADORA

Prof. Msc. Ricardo Jardel Nunes da Silveira
(Orientador)
Universidade Federal do Ceará (UFC)

Prof. Msc. Marcelo Araújo Lima
(Co-Orientador)
Instituto Federal do Ceará (IFCE)

Prof. Dr. Jarbas Aryel da Silveira
Universidade Federal do Ceará (UFC)

Prof. Msc. Daniel Alencar Barros Tavares
Instituto Federal do Ceará (IFCE)

Dedico este trabalho à minha família e namorada, pessoas que fizeram de tudo para que eu chegasse onde cheguei.

Agradecimentos

Agradeço primeiramente a Deus, que iluminou meu caminho durante essa jornada, me dando saúde e força para superar as dificuldades.

À minha namorada, Luéline Elias, pelo amor, paciência, dedicação e companheirismo em todos os momentos.

À minha família, por sua capacidade de acreditar e investir em mim. Mãe, sua dedicação foi o que deu, em alguns momentos, a esperança para seguir.

Ao meu orientador, Prof. Ricardo Jardel Nunes da Silveira, pelo acompanhamento e estreitamento da relação professor-aluno e exemplo de profissional bem como pelo apoio, incentivo, sugestões e comentários durante a supervisão dos meus estudos.

Ao meu coorientador, Prof. Marcelo Araújo Lima, pelo apoio, incentivo, sugestões e tempo dedicado para me ajudar durante meus estudos.

Aos meus amigos da Universidade Federal do Ceará, 8086FC e 8086Team pela amizade e pelos momentos de descontração e estudo.

E a todos que direta ou indiretamente fizeram parte da minha formação, o meu muito obrigado.

"A persistência é o caminho do êxito."
(Charles Chaplin)

Resumo

Palavras-chaves: Detecção de ataques DDoS. Segurança em redes. Tempo real. Framework .

Abstract

Key-words: DDoS attack detection. Network Security. Real-time. Framework.

Lista de ilustrações

Lista de tabelas

Lista de abreviaturas e siglas

DDOS	Distributed Denial of Service

Lista de símbolos

X	Vetor de entrada para correlação NaHiD

Sumário

1 Introdução

Ataques Distribuídos de Negação de Serviço (do inglês, DDoS) são uma ameaça a servidores de redes online, tais como servidores de sites web e servidores em nuvem. O objetivo desse tipo de ataque intencional é inundar o alvo com requisições e assim deixá-lo indisponível na rede. Existem essencialmente três tipos de ataques: Negação distribuída, Handshake e UDP. O primeiro caracteriza-se por requisições abertas por um grande número de computadores infectados. No segundo, faz-se uma comunicação inicial com o alvo que não é completada, mantendo assim o servidor esperando indefinidamente. Já no terceiro, fluxos falsos UDP são criados com o mesmo objetivo de tornar o serviço inoperante. Os métodos estatísticos existentes na literatura para análise de ataque DDoS falham principalmente devido às correlações de deslocamento, escala e deslocamento-escala ao longo de tráfegos de rede, gerando assim uma grande ocorrência de falsos positivos. Além disso, métodos estatísticos impõem alto overhead computacional quando um grande número de objetos é incluído para análise. Consequentemente, tais métodos falham em realizar detecção de ataque DDoS em tempo real. Algumas medidas de correlação tais como Pearson, Spearman e Kendall falham em identificar a diferença entre um pacote normal e um malicioso quando há valores correlacionados entre os pacotes. De fato, um método de detecção de ataques DDoS precisa considerar poucos parâmetros de tráfego durante a análise, tal como o método chamado NaHiDVERC (HOQUE; KASHYAP; BHATTACHARYYA, 2017), o qual analisa apenas entropia de IPs e taxa de pacote. Tendo em vista uma implementação em software e hardware, este método será utilizado em nosso trabalho, visto que é facilmente implementável em hardware. O método computa dois valores: a distância absoluta e o desvio entre A e B a partir da média e do desvio padrão. Se a entropia de IPs origem em um pequeno intervalo de tempo é alta e a taxa de pacote é também muito alta, a probabilidade de ataque é alta. Se a variação entre IPs origem é muito alta e a taxa de pacote também é alta, a probabilidade de ataque é alta. O framework tem como objetivo detectar ataques DDoS em tempo real no computador alvo. Trata-se de uma combinação entre aplicações em software e hardware, para classificar um tráfego como normal ou ataque com uma taxa aceitável de acertos. Tal arcabouço possui três componentes: pré-processamento, um módulo de hardware dedicado para detecção e um gerente de segurança. Neste trabalho os componentes um e três serão trabalhados. Além disso, é necessária a presença de um roteador para capturar tráfego e duas bases de dados. Amostras de tráfego serão capturadas de uma porta do roteador como um pacote TCP/IP, que são enviadas ao módulo de pré-processamento. Nessa fase, a cada segundo, os pacotes recebidos são agrupados e essa instância de tráfego é enviada para o módulo de detecção de ataque, que irá classificar a instância como normal ou ataque. O gerente

de segurança manterá um perfil normal e um valor limiar em sua base de perfis, para ser usado pelo módulo de detecção. Incrementalmente, o gerente recalcula o perfil normal e o limiar baseado nos valores anteriores. Existem duas abordagens durante a análise do tráfego: uma considerando apenas a informação no cabeçalho do pacote ou se o cabeçalho e dados estarão juntos. Nas duas formas os campos dos pacotes são analisados para detectar alguma anomalia na rede. IP e porta origem/destino, protocolos e flags do cabeçalho TCP são úteis para detectar pacotes maliciosos. Assim, a entropia e a variação entre IPs origem e taxa de pacotes são calculados para cada amostra de tráfego. O último módulo, que é o módulo de segurança irá operar offline e fará análises detalhadas dos logs de detecção usando técnicas de machine learning e estatística. Além disso, feedbacks de especialistas podem ser utilizados para validar os resultados. Inicialmente, o gerente vai calcular um perfil de tráfego normal que melhor representa instâncias desse tráfego para treinamento. Esses valores serão carregados na base de dados. Vale ressaltar que esses valores serão modificados dinamicamente de acordo com as amostras de tráfego.

1.1 Objetivos

1.2 Organização da monografia

Os estudos deste trabalho estão organizados da seguinte forma: No próximo capítulo será apresentado um estudo bibliográfico sobre ameaças de rede e ataques DDoS. No Capítulo 3, a modelagem do ambiente de simulação utilizado neste trabalho é descrita. No Capítulo 4, apresentamos o desempenho obtido pelo *framework* estudado por meio da taxa de acerto para cada janela de tráfego. Por fim, o último capítulo deste trabalho apresenta as conclusões realizadas a partir dos resultados obtidos e algumas perspectivas para a continuação deste trabalho.

2 Revisão Bibliográfica

Falar sobre ataques, definir objeto de tráfego, sniffer ...

3 Metodologia

Nesse capítulo são apresentadas a medida de correlação utilizada no trabalho, além das principais características do *framework*, mostrando como a correlação é aplicada para a detecção de ataques DDoS e quais bases de dados são utilizadas para a avaliação do *framework*, destacando sua estrutura e ferramentas utilizadas para o tratamento dos dados.

3.1 Modelo de correlação NaHiD

Neste trabalho, o *framework* utilizado baseia-se na correlação proposta por (HOQUE; KASHYAP; BHATTACHARYYA, 2017) chamada NaHiD (nome que possivelmente provém a partir das iniciais de cada autor), cujo objetivo é distinguir objetos de tráfego normais e maliciosos. Tal medida leva em consideração principalmente o desvio padrão e a média de cada objeto, ponderando cada elemento como mostrado na equação a seguir:

$$NaHiD(X,Y) = 1 - \frac{1}{n}\sum_{i=1}^{n}\frac{(|X(i) - Y(i)|)}{||\mu X - sX| - X(i)| + ||\mu Y - sY| - Y(i)|} \tag{3.1}$$

onde

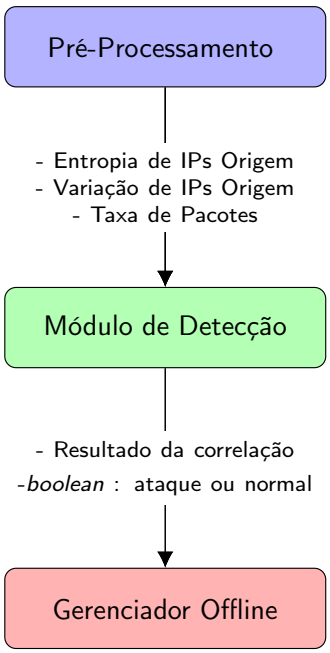
- μX: Média aritmética do objeto de tráfego X.
- μX: Média aritmética do objeto de tráfego Y.
- sX: Desvio padrão do objeto de tráfego X.
- sY: Desvio Padrão do objeto de tráfego Y.

As provas de simetria e identidade da correlação podem ser encontradas em (HOQUE; KASHYAP; BHATTACHARYYA, 2017).

3.2 Framework de detecção de ataques DDoS

O *framework* tem como objetivo, detectar ataques DDoS em tempo real na rede monitorada, a partir de dados trafegados na rede com uma taxa aceitável de erros. Tal arcabouço possui três módulos: pré-processamento, detecção e um de segurança. A Figura 3.1 mostra o fluxo de funcionamento do *framework*. Amostras de tráfego são capturadas de uma porta do roteador na forma de um pacote TCP/IP e enviadas ao módulo de pré-processamento. Nessa fase, a cada segundo, os pacotes recebidos são agrupados e

Figura 3.1 – Estrutura do *framework* analisado



Fonte: Elaborada pelo autor.

essa instância de tráfego é enviada para o módulo de detecção de ataques, que irá classificar a instância como normal ou maliciosa. O gerente de segurança manterá um perfil normal, como referência, e um valor limiar de correlação em sua base de perfis, para ser usado pelo módulo de detecção. Incrementalmente, o gerente recalcula o perfil normal baseado nos valores anteriores.

3.2.1 Pré-Processamento

Nessa etapa, os dados são coletados por um *sniffer* da rede, o qual analisa todos os pacotes trafegados e, a cada segundo, as métricas desejadas são calculadas para servirem de entrada para a correlação NaHiD.

3.2.1.1 Entropia de IPs origem

A entropia de IPs origem é uma medida do grau de desordem, onde ela é máxima caso todos os elementos sejam diferentes e o tamanho da entrada seja máximo, e será mínima (igual a 0) quando todos os elementos forem iguais, independentemente do tamanho. Assim, a entropia é dada pela seguinte fórmula:

$$H(X) = -\sum_{i}^{n} p(x_i) log_2(x_i) \tag{3.2}$$

Onde X é a entrada e representa os IPs origem das requisições e n é o número total de valores possíveis para o IP origem. A Tabela 1 mostra exemplos com valores de entrada

para entropia, bem como o resultado do cálculo da função.

Note que a entropia é mínima quando todos os IPs origem são iguais (primeira linha da

Tabela 1 – Exemplo de IPs origem com respectivos valores de entropia

IPs origem					Entropia
192.168.8.8	192.168.8.8	192.168.8.8	192.168.8.8	192.168.8.8	0
192.168.15.129	192.168.8.5	192.168.8.8	192.168.10.16	192.168.20.22	2.3219
192.168.8.8	192.168.8.8	192.168.8.5	192.168.10.16	192.168.20.22	1.9219
192.168.20.22	192.168.20.22	192.168.20.22	192.168.20.22	192.168.8.8	0.7219

Fonte: Elaborada pelo autor.

tabela) e máxima quando todos os os IPs origem são diferentes (segunda linha).

3.2.1.2 Variação de IPs Origem

Essa medida, diferentemente da entropia, trata-se da taxa de mudança dos IPs origem e é calculada da seguinte forma:

$$V_{Ip}(X) = \frac{\delta}{N} \tag{3.3}$$

Onde δ é o número de mudanças de IPs origem e N é o numero total de IPs de entrada. Neste trabalho consideramos uma variação cada troca de valores como no exemplo:

$$X = 1, 2, 1, 2, 3 \tag{3.4}$$

Assim, nesse vetor consideram-se 4 variações ainda que sejam para um valor que repetiu-se. Assim se os IPs origem mudarem frequentemente, a variação será alta. (HOQUE; KASHYAP; BHATTACHARYYA, 2017)

A observação do comportamento de ataques por *flood* mostra que esse tipo de ameaça pode ser gerada por atacantes reais como zumbis. Se endereços de IP origem falsificados forem utilizados durante um ataque DDoS TCP SYN, a entropia e variação de IPs origem serão altas e esse comportamento também ocorre em um tráfego normal. (HOQUE; KASHYAP; BHATTACHARYYA, 2017). Assim faz-se necessário o uso da taxa de pacotes em bits como terceiro parâmetro de entrada para o módulo de detecção.

3.2.2 Módulo de Detecção

O modulo de detecção consiste na aplicação da correlação NaHiD, utilizando os três parâmetros de entrada fornecidos pelo módulo de pré-processamento:

- Variação de IPs origem
- Entropia de IPs origem

- Taxa de pacotes

Por tratar-se de uma medida de correlação é necessário manter um valor de referência para o cálculo. Assim, os parâmetros(Variação e entropia de IPs origem, além da taxa de pacotes) de um tráfego normal devem ser fixados para a comparação com a instância de tráfego a ser analisada. Além disso, define-se um limiar do resultado da correlação para distinguir instâncias normais de maliciosas. Caso a correlação calculada seja menor que esse limiar, tal tráfego em analisado não tem semelhança suficiente com o perfil normal comparado, sendo considerado um ataque.

3.2.3 Gerenciador Offline

Nessa módulo, os valores de correlação, IPs origem, destino, taxas de pacotes, além dos resultados do módulo de detecção são salvos para a janela de tráfego em análise e se o módulo de detecção identificar que o tráfego em questão é normal, este será atualizado com os valores do mesmo para a próxima janela.

3.3 Aplicação do *framework* de detecção em bases de dados reais

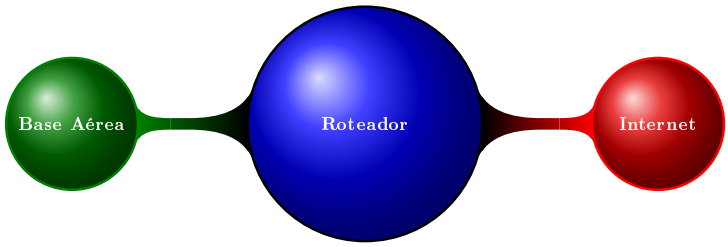
Para a avaliação do trabalho, duas bases de tráfegos de rede foram escolhidas: DARPA e DataMining[escolher melhor esse nome] os quais são mais detalhados a seguir

3.3.1 DARPA - MIT

A base de dados DARPA foi produzida por pesquisadores do *Lincoln Laboratory* do Instituto de Tecnologia de Massachusetts nos Estados Unidos e tem por objetivo coletar dados de tráfego de rede da Força Aérea do país para encontrar vulnerabilidades em seu sistema bem como ser utilizado para avaliações futuras. Os dados foram coletados e passaram por uma fase de treinamento de 7 semanas com 38 tipos de ataques para simular ameaças internas a rede. O ambiente de rede era composto por duas partes: a rede interna da Força aérea e a rede externa que representava a Internet; ambos conectados por meio de um roteador como mostra a Figura 3.2. Assim, um *sniffer* de rede foi instalado no roteador e todas as requisições para os computadores da Força aérea foram capturadas em um arquivo tcpdump, o qual pode ser encontrado em (DARPA...,).
A partir desse *dataset* é possível extrair informações acerca de cada pacote transmitido durante o período de aquisição dos dados como mostra o exemplo na Tabela 2.

No presente trabalho ferramentas como edicap e tcpdump foram utilizadas para o tratamento desse *dataset* no módulo de processamento. Assim, algumas considerações devem ser feitas:

Figura 3.2 – Estrutura de rede Base Aérea dos EUA



Fonte: Elaborada pelo autor.

Tabela 2 – Exemplo base de dados DARPA

Número	**Tempo**	**Origem**	**Destino**	**Protocolo**	**Tamanho**[bytes]
1	18:56:12.1386	192.168.0.20	192.168.0.30	TCP	60
2	18:56:12.1391	192.168.0.30	192.168.0.20	TCP	60
3	18:56:12.1588	192.168.0.30	192.168.0.20	TELNET	84
4	18:56:12.2099	192.168.0.20	192.168.0.30	TCP	60
5	18:56:13.0567	192.168.0.20	192.168.0.30	TELNET	69
6	18:56:13.0584	192.168.0.30	192.168.0.20	TELNET	66
7	18:56:13.0626	192.168.0.20	192.168.0.30	TELNET	72
8	18:56:13.0821	192.168.0.30	192.168.0.20	TCP	60

Fonte: Elaborada pelo autor, baseada em (LIPPMANN *et al.*, 2000).

- Janela de um segundo de tráfego.
- Cálculo de entropia, variação de IPs origem e taxa de pacotes média.
- Cálculo da correlação NaHiD com base no item anterior.

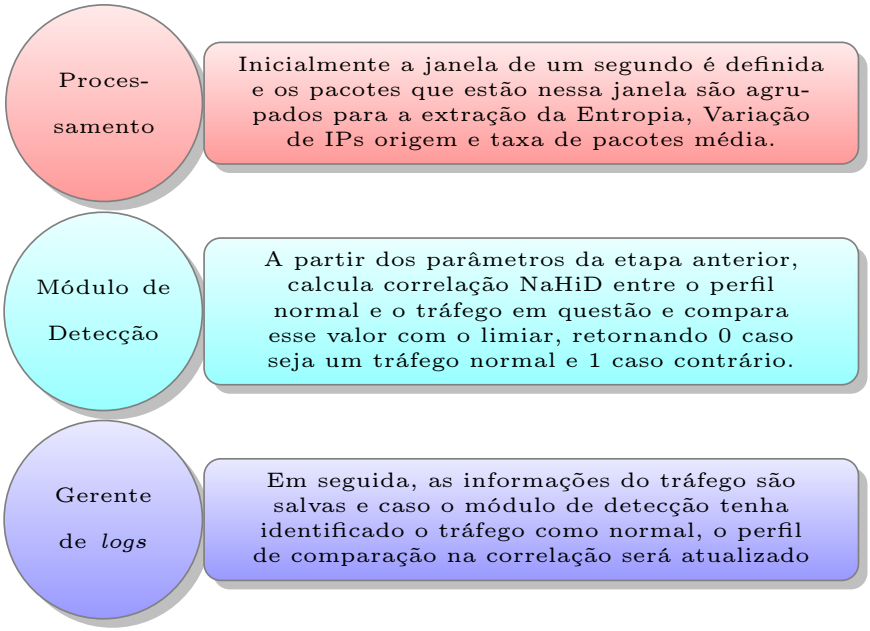
Note que de acordo com a estrutura do *dataset* mostrada na Tabela 2, a implementação do *framework* segue o seguinte fluxo mostrado na Figura 3.3

.

3.3.2 DataMining

Outra base de dados estudada no trabalho foi a desenvolvida por (ALKASASSBEH *et al.*, 2016) a qual consta em sua totalidade por ataques DDoS de quatro tipos:

Figura 3.3 – Diagrama descritivo análise e aplicação do *framework* na base de dados DARPA



Fonte: Elaborada pelo autor.

- SIDDoS
- HTTP Flood
- UDDP Flood
- Smurf

A Tabela 3 mostra os campos do *dataset*

Algumas considerações foram tomadas para a análise dessa base de dados:

- Para construir a janela de um segundo, considerou-se a soma de todos os atrasos por pacote:
  - Atraso de nó do pacote.
  - Atraso de pacote.
  - Tempo de pacote reservado.
- A média das taxas dos pacotes foi considerada dentro da janela de um segundo.
- Por ser um *dataset* composto apenas por ataques, a comparação com o limiar inverte-se para denotar o quanto dois pacotes são parecidos na correlação.

A base de dados é disponibilizada no formato *Weka Attribute-relation*(extensão arff), o qual é utilizado geralmente para compactar grandes massas de dados e processá-las

Tabela 3 – Estrutura base de dados (ALKASASSBEH *et al.*, 2016)

Número	**Tempo**
1	Endereço IP origem
2	Endereço IP destino
3	Id do pacote
4	Nó origem
5	Nó destino
6	Tipo de pacote
7	Tamanho do pacote
8	Flags
9	Id da flag
10	Número de sequência
11	Número de pacotes
12	Número de bytes
13	Nome do nó origem
14	Nome do nó destino
15	Entrada de pacote
16	Saída de pacote
17	Taxa de pacotes Recebidos
18	Atraso de nó do pacote
19	Taxa de pacotes
20	Taxa de bytes
21	Tamanho médio do pacote
22	Utilização
23	Atraso de pacote
24	Tempo de envio do pacote
25	Tempo de pacote reservado
26	Primeiro pacote enviado
27	Último pacote reservado

Fonte: Elaborada pelo autor, baseada em (ALKASASSBEH *et al.*, 2016).

utilizando técnicas de *machine learning*. Assim, para o processamento dos mesmos as ferramentas Weka e MATLAB foram utilizadas.

4 Resultados

5 Conclusões e Trabalhos Futuros

Referências

ALKASASSBEH, M. *et al.* Detecting Distributed Denial of Service Attacks Using Data Mining Techniques. v. 7, 01 2016. Citado 3 vezes nas páginas 9, 20 e 22.

DARPA INTRUSION DETECTION EVALUATION. <https://www.ll.mit.edu/ideval/index.html>. Acessado em 18/08/2017.

Citado na página 19.

HOQUE, N. *et al.* Real-time DDoS Attack Detection Using FPGA. **Computer Communications**, v. 110, n. Supplement C, p. 48 – 58, 2017. ISSN 0140-3664. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0140366416306442>. Citado 3 vezes nas páginas 13, 16 e 18.

LIPPMANN, R. P. *et al.* Evaluating intrusion detection systems: the 1998 darpa off-line intrusion detection evaluation. In: **DARPA Information Survivability Conference and Exposition, 2000. DISCEX '00. Proceedings**. [S.l.: s.n.], 2000. v. 2, p. 12–26 vol.2.

Citado na página 20.