UDESC UNIVERSIDADE DO ESTADO DE SANTA CATARINA

Seminário de Iniciação Científica
Universidade do Estado de Santa Catarina
29º SIC UDESC

Plataforma para Análise de Dados de Honeypots no Curto Prazo

Gustavo José Neves da Silva[1], Rafael R. Obelheiro[2]

[1] Acadêmico(a) do Curso de Ciência da Computção CCT - bolsista PROBIC/UDESC
[2] Orientador, Departamento de Ciência da Computação – rafael.obelheiro@udesc.br

Palavras-chave: Redes de computadores, Honeypot, DDoS

Honeypots são ferramentas utilizadas a fim de realizar a coleta e observação do tráfego de rede relativo a ataques [1]. Os arquivos de dados produzidos por *honeypots* muitas vezes são processados e analisados manualmente. Isso pode fazer com que anomalias e mudanças de tendências no tráfego sejam percebidas muito tempo após terem acontecido, impossibilitando uma investigação mais aprofundada sobre essas ocorrências e suas causas. Para contornar esse problema, é possível automatizar algumas análises de dados de *honeypots*. O presente trabalho tem como objetivo propor uma plataforma para o processamento automatizado diário dos dados coletados por *honeypots*. Essa plataforma será usada no processamento de dados do HReflector [1], um *honeypot* para a observação de ataques distribuídos de negação de serviço por reflexão (DRDoS, *distributed reflection denial of service*), a fim de gerar estatísticas e gráficos (tráfego por protocolo, ASN e país) relativos ao tráfego diário.
O HReflector usa o Tcpdump [2] para captura de tráfego, e armazena os dados em *chunks* de 100 MB, no formato PCAP [2, 3]. Com isso, o tráfego de um dia pode estar pulverizado em vários arquivos. O fluxo de processamento da plataforma proposta é mostrado na Figura 1. No passo 1, determina-se quais arquivos PCAP contêm o tráfego do dia de interesse, usando a ferramenta `capinfos` [3]. No passo 2, esses arquivos são combinados em um único arquivo usando a ferramenta `mergecap` [3]. No passo 3, os dados de pacotes contidos no arquivo PCAP são agregados em fluxos (*flows*), ou seja, dados de sessões em intervalos de 5 minutos, usando a ferramenta `argus` [4]. No passo 4, o arquivo de fluxos é convertido para um arquivo de texto puro usando a ferramenta `ra` [4]. No passo 5, o arquivo gerado no passo 4 é processado para obter tabelas e gráficos com estatísticas de tráfego por ASN (5a), protocolo (5b) e país (5c), usando o pacote estatístico R [5]. Todo o fluxo de processamento é orquestrado por um script R. A Figura 2 mostra uma tabela e um gráfico gerados pela ferramenta, com estatísticas do tráfego diário por país. Pode-se observar nesse exemplo que, no dia 29/09/2018, a maior parte das vítimas de ataques DRDoS estava situada nos Estados Unidos da América(US), com o total de 32.61 MB, e que o pico de tráfego ocorreu às 01:00, com 1.71 KB/s. O ataque com maior

agressividade quanto a velocidade, teve vítimas situadas na República Popular da China, com pico de tráfego às 22:35, com 22.29KB/s.

Embora desenvolvida com foco no HReflector, a plataforma pode ser facilmente adaptada a outros *honeypots* que armazenem dados nos formatos PCAP ou de fluxos. Na continuidade desta pesquisa, pretende-se colocar a plataforma em produção e disponibilizar seu código fonte sob uma licença livre.

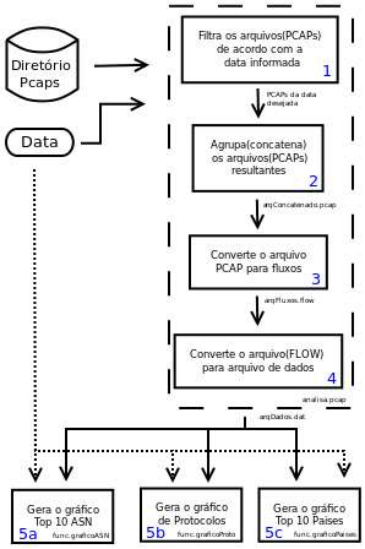
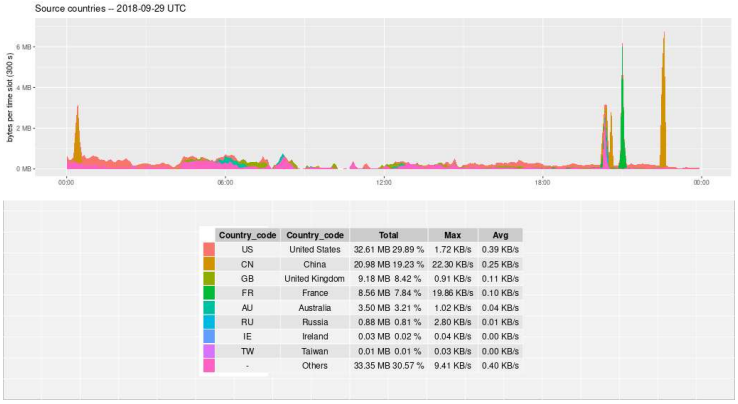


Fig. 1 *Diagrama*

Fig. 2 *Gráfico Países*



Country_code	Country_code	Total	Max	Avg
US	United States	32.61 MB 29.89 %	1.72 KB/s	0.39 KB/s
CN	China	20.98 MB 19.23 %	22.30 KB/s	0.25 KB/s
GB	United Kingdom	9.18 MB 8.42 %	0.91 KB/s	0.11 KB/s
FR	France	8.56 MB 7.84 %	19.86 KB/s	0.10 KB/s
AU	Australia	3.50 MB 3.21 %	1.02 KB/s	0.04 KB/s
RU	Russia	0.88 MB 0.81 %	2.80 KB/s	0.01 KB/s
IE	Ireland	0.03 MB 0.02 %	0.04 KB/s	0.00 KB/s
TW	Taiwan	0.01 MB 0.01 %	0.03 KB/s	0.00 KB/s
-	Others	33.35 MB 30.57 %	9.41 KB/s	0.40 KB/s

Referências:
1. HEINRICH, T. Caracterização de Ataques DRDoS usando Honeypot. Dissertação de Mestrado em Computação Aplicada, Universidade do Estado de Santa Catarina, 2019.
2. TCPDUMP/LIBPCAP Public Repository. URL http://www.tcpdump.org/
3. Wireshark. URL https://www.wireshark.org/
4. ARGUS - Auditing Network Activiy. URL https://www.qosient.com/argus/
4. R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/