UNIVERSIDADE FEDERAL DO AMAZONAS

INSTITUTO DE CIÊNCIAS EXATAS

DEPARTAMENTO DE CIÊNCIA DA COMPUTAÇÃO

PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA

DETECÇÃO DE ANOMALIAS NA INTERNET ATRAVÉS DA ANÁLISE DO TRÁFEGO DNS

KAIO RAFAEL DE SOUZA BARBOSA

MANAUS
MAIO 2010

Livros Grátis

http://www.livrosgratis.com.br

Milhares de livros grátis para download.

UNIVERSIDADE FEDERAL DO AMAZONAS

INSTITUTO DE CIÊNCIAS EXATAS

DEPARTAMENTO DE CIÊNCIA DA COMPUTAÇÃO

PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA

KAIO RAFAEL DE SOUZA BARBOSA

DETECÇÃO DE ANOMALIAS NA INTERNET ATRAVÉS DA ANÁLISE DO TRÁFEGO DNS

Dissertação apresentada ao Programa de Pós-Graduação em Informática da Universidade Federal do Amazonas, como requisito parcial para obtenção do título de Mestre em Informática, linha de pesquisa Redes de Computadores.

Orientador: Prof. Dr. Eduardo James Pereira Souto

MANAUS

MAIO 2010

Ficha Catalográfica
(Catalogação realizada pela Biblioteca Central da UFAM)

KAIO RAFAEL DE SOUZA BARBOSA

DETECÇÃO DE ANOMALIAS NA INTERNET ATRAVÉS DA ANÁLISE DO TRÁFEGO DNS

Dissertação apresentada ao Programa de Pós-Graduação em Informática da Universidade Federal do Amazonas, como requisito parcial para obtenção do título de Mestre em Informática, linha de pesquisa Redes de Computadores.

Aprovado em __ de Maio de 2010

Banca Examinadora

_____

Prof. Dr. Eduardo James Pereira Souto (Orientador)
Universidade Federal do Amazonas

_____

Prof. Dr.-Ing. Edjair de Souza Mota
Universidade Federal do Amazonas

_____

Prof. Dr. Djamel F. H. Sadok
Universidade Federal de Pernambuco

Agradecimentos

Senhor meu Deus, tenho muito a lhe agradecer por ter me dado forças em todos os momentos mais difíceis durante essa fase da minha vida. Obrigado por tudo, Senhor Jesus Cristo.

Minha Família! Peço perdão por não estar presente em diversas situações em que fui necessário. Esse trabalho é uma homenagem ao teu amor e dedicação. Muito obrigado.

Diana Guimarães, meu amor, linda, se não fosse o teu companheirismo, amor e amizade, minha caminhada até aqui seria muito difícil. Obrigado pela paciência, dedico o meu sucesso à você. Amo-te por tudo.

Meus amigos, João "minu" Luis, Rafael Sousa, Regeane Aguiar, João Batista, Rick "Parente", Daniel Costa, Saulo Queiroz, Andrea Giordanna, Arlen Nascimento, Christophe Xavier, Marcela Pessoa, Efren Lopes, Edna Magalhães, Sionise Rocha, Vandermi "maninho" Silva, Lady Daiana e Viviane Gomes, foi um enorme prazer compartilhar conhecimentos, aflições, sorrisos, dicas e truques no processo de construção do saber durante o mestrado. Obrigado.

Professor Eduardo Feitosa, José Pinheiro, Jansen Sena, Raimundo Barreto, José Francisco, Horácio Fernandes, Edjair Mota, Marco Cristo, Eduardo Nakamura e Leandro Galvão. Muito obrigado pelos conselhos, aulas e conversas no corredor.

Meu orientador, Professor Eduardo Souto, tenho muito a lhe agradecer pela credibilidade, confiança no meu trabalho, paciência e, acima de tudo, humildade e competência técnica para gerir projetos críticos e pessoas. Muito obrigado pelos comentários e puxões de orelha. Sem eles, não teria aprendido tanto durante o processo de pesquisa.

Obrigado Elienai Nogueira pela sua ajuda durante o mestrado. A secretaria sempre exercerá um papel importantíssimo no DCC. Muito obrigado.

À CAPES, pelo apoio financeiro. Obrigado.

Ao projeto da OARC pela parceria com a Universidade Federal do Amazonas. Em especial, ao Duane Wessels, pelas dicas e ajuda durante a fase de codificação deste trabalho.

E finalmente, todos aqueles que compartilharam e contribuíram de alguma forma para realização do Mestrado.

Lista de Ilustrações

Lista de Abreviações e Siglas

DNS	*Domain Name System*
DoS	*Denial of Service*
DDoS	*Distributed Denial of Service*
IP	*Internet Protocol*
SSH	*Secure Shell*
SMTP	*Simple Mail Transfer Protocol*
TTL	*Time to Live*
SGBD	Sistema Gestor de Base de Dados
TCP	*Transmission Control Protocol*
UDP	*User Datagram Protocol*

Lista de Tabelas

Sumário

Abstract

Today, most Internet service is based on a model of operation where some query to the DNS system is performed before any communication activity. Analyzing the DNS traffic is expected to find typical behaviors, such as requests that use the default class of queries and some well-defined resource records. However, analyzing the message protocol, it is also possible to identify unwanted traffic or anomalous, i.e traffic that should not be present in the DNS queries or responses.

Due to the importance of DNS, this work proposes and develops a methodology using entropy for detection, identification and classification of anomalies through DNS traffic. This method is based on the correlation of information in the IP header (such as addresses and ports of origin) and the DNS protocol to infer network anomaly behavior.

The goal of this work is to use the concept of entropy to distinguish between normal and anomalous behavior in network traffic. To better understand the database, an overview is presenting of network anomalies from the analysis of passive data. Then, the entropy concept is demonstrated. The results show that the proposed approach is able to identify changes in the distribution of traffic and demonstrates its ability to characterize classes of anomalous behavior from validation of records of real traffic network.

The main contribution of this work is the utilization of DNS traffic combined with information from the IP header for the detection, identification and classification of network anomaly.

Keywords: Networks, Network Security, Network Anomaly, DNS.

Resumo

Atualmente, a maioria dos serviços da Internet é baseada em um modelo de funcionamento em que alguma consulta ao sistema DNS é realizada antes de qualquer atividade de comunicação. Através da análise do tráfego DNS é esperado encontrar comportamentos típicos, como solicitações que utilizam a classe padrão de consultas e alguns registros de recursos bem definidos. Entretanto, por meio da análise da mensagem do protocolo, também é possível identificar tráfego não desejado ou anômalo, ou seja, tráfego que não deveria estar presente nas consultas ou respostas do DNS.

Devido à importância do DNS, este trabalho propõe e desenvolve uma metodologia utilizando entropia para detecção, identificação e classificação de anomalias através do tráfego DNS. Este método é baseado na correlação de informações presente no cabeçalho IP (como endereços e portas de origem) e no protocolo DNS para inferir sobre anomalias de rede.

O objetivo deste trabalho é a utilização do conceito de entropia para distinguir entre o comportamento normal e anormal no tráfego de rede. Para o entendimento da base de dados é apresentando uma visão geral das anomalias de rede a partir da análise passiva dos dados. Em seguida, o conceito de entropia é demonstrado. Os resultados mostram que a abordagem proposta é capaz de analisar as mudanças na distribuição de recursos de tráfego e demonstra a sua capacidade para caracterizar classes de comportamento anômalo a partir de validações de registros de tráfego real de rede.

A principal contribuição deste trabalho é a utilização dos recursos de tráfego DNS combinados com informações do cabeçalho IP para detecção, identificação e classificação de anomalias de rede.

Palavras-chave: Redes, Segurança de Redes, Anomalias de Rede, DNS.

1 Introdução

"O futuro dependerá daquilo que fazemos no presente."

Mahatma Gandhi

1.1 Contextualização

O protocolo DNS (*Domain Name System*) [1] possui um papel crucial para o funcionamento da Internet: a tradução de nomes de máquinas em endereços IP (*Internet Protocol*). Tal mecanismo de tradução permite que as aplicações possam obter informações sobre os mais variados tipos serviços disponíveis na rede como, por exemplo, a localização de servidores de correio eletrônico, servidores web, aplicações de comércio eletrônico (*e-commerce*), ou ainda, outros servidores de nomes.

Atualmente, a maioria dos serviços da Internet é baseada em um modelo de funcionamento em que alguma consulta ao sistema de resolução de nomes é realizada antes de qualquer atividade de comunicação. Nesse contexto, técnicas de medição e a análise de informações de tráfego DNS permitem inferir e melhor compreender as particularidades do comportamento da rede. Portanto, as mensagens DNS oferecem informações que identificam comportamentos típicos, como solicitações empregando classe padrão de consultas e registros de recursos bem definidos. Entretanto, investigações do tráfego DNS também detectam comportamentos não desejados ou anômalos, isto é, tráfego que não deveria estar presente nas consultas ou respostas do DNS.

As causas dessas anomalias no tráfego incluem, por exemplo, perguntas com nomes de domínios de primeiro nível inválidos, solicitações geradas a partir de servidores DNS com problemas de configuração em zona de domínio, ataques de reconhecimento de rede [2], exército de máquinas controladas remotamente por criminosos (*botnets*) [3] ou máquinas

infectadas por aplicações de código malicioso como cavalos de tróia (*trojans*) e vermes (*worms*) [4].

Na prática, o a análise do tráfego DNS pode ser útil no projeto e planejamento de redes, na engenharia de tráfego, na qualidade de serviço e gerenciamento de redes. A análise de padrões de tráfego DNS pode ser usada como ferramenta para a detecção, identificação e quantificação de anomalias de rede, independente de terem sido causadas intencionalmente ou não. O diagnóstico de anomalias apresenta grandes desafios, pois é necessário extrair padrões anômalos de grandes volumes de dados e, como citado, as causas das anomalias podem ser bastante variadas [5].

Este trabalho apresenta uma nova metodologia para detecção de anomalias baseada em medidas de entropia [6] a partir da análise do tráfego DNS. Este método utiliza a correlação de informações presente no cabeçalho IP (como endereços e portas de origem) e no protocolo DNS para inferir sobre anomalias de rede.

1.2 Motivação

Uma breve análise do tráfego na Internet comprova o crescente aumento no transporte do tráfego considerado não solicitado, improdutivo e muitas vezes ilegítimo. Originado através de atividades como, por exemplo, mensagens eletrônicas não solicitadas (SPAM); atividades fraudulentas como *phishing*[1] e *pharming*[2]; proliferação de vírus e vermes, entre outros. Uma característica comum a todas essas anomalias é que elas, tipicamente, usam o protocolo DNS para encontrar suas vítimas.

Em razão das funcionalidades e vulnerabilidades presente no sistema DNS e o constante avanço das técnicas de ataques, é importante que outros tipos de abordagens relacionadas à segurança, para esse serviço, sejam investigadas e propostas. A análise do tráfego DNS é um

[1] *Phishing* é um tipo de fraude eletrônica caracterizada pela tentativa de obter informações pessoais privilegiadas (por exemplo, números de cartões de créditos e senhas) através de sites falsos ou mensagens eletrônicas forjadas.

[2] *Pharming* é o termo atribuído ao ataque baseado no envenenamento de cache DNS que consiste em corromper o DNS em uma rede de computadores, fazendo com que a URL de um *site* passe a apontar para um servidor diferente do original.

ponto de partida para a descoberta e interceptação precoce do tráfego gerado por anomalias de rede.

Neste contexto, é preciso identificar características que não seguem os padrões de comportamento de comunicação na rede, como ataque de negação de serviço ou falhas de configuração no serviço do DNS. Essas anomalias resultam em sobrecarga nos servidores de nome raiz (*Root Servers*) da Internet, ou também, nova atividades maliciosas que ainda não são publicamente conhecidas (*zero-day attacks*) que degradam os recursos de infraestrutura da Internet.

Para detectar padrões de comportamento anômalo no tráfego de rede, algumas abordagens [7] [8] têm utilizado entropia para indicar o grau de concentração ou dispersão de uma distribuição de probabilidade. Em outras palavras, a entropia fornece um sumário do comportamento do tráfego indicando alterações significativas nos padrões de comunicação na rede.

Técnicas de medição da teoria da informação, como a entropia, oferecem características importantes para detecção de anomalias na Internet como a combinação e visualização de fluxo da quíntupla já conhecida: endereço IP de origem, endereço IP de destino, porta de origem, porta de destino e protocolo de rede. A teoria da informação é importante para extração e classificação desses fluxos, por isso, a entropia possui ênfase na detecção de anomalias.

Por outro lado, essas soluções tomam como base recursos do tráfego bem conhecidos para detectar tráfego malicioso na rede. Enquanto que, anomalias que usam o tráfego DNS para seqüestrar, interromper ou atacar outros nós de rede estão localizadas na mensagem DNS. Logo, para identificar tráfego anômalo através da Entropia, é necessário também correlacionar campos da mensagem DNS.

1.3 Objetivos

O objetivo geral deste trabalho é propor e demonstrar uma metodologia para detecção, identificação e caracterização de anomalias na Internet através da análise do tráfego DNS.

Este trabalho utiliza conceitos da Teoria da Informação, especificamente o conceito de Entropia, como métrica para sinalizar comportamentos maliciosos no tráfego da Internet. O

objetivo é caracterizar o comportamento de anomalias de rede que utilizam o protocolo DNS. Para isso, será necessário definir quais informações do protocolo DNS devem ser usadas para detectar anomalias de rede. Por último, uma avaliação é realizada a partir da implementação de um protótipo da metodologia proposta usando o tráfego real da Internet.

1.4 Principais contribuições

A principal contribuição deste trabalho está na utilização da Entropia para detecção de anomalias através do tráfego DNS. Enquanto outros trabalhos detectam comportamentos maliciosos por meio da tupla bem conhecida, este trabalho sugere uma nova metodologia que correlaciona alguns campos do cabeçalho IP e informações da mensagem DNS para detectar, identificar e classificar anomalias de rede.

O emprego da Entropia no processo de detecção de anomalias é um ponto positivo, pois essa métrica fornece um sumário dos comportamentos mais significativos do tráfego, isso significa que, quando aplicada em tráfego de redes de alta velocidade, não será necessário observar cada evento de rede, apenas os mais relevantes. Em outras palavras, apenas o tráfego que está fora do padrão de rede.

A metodologia proposta utiliza alguns conceitos definidos em Xu [8] e os emprega no contexto de anomalias através do tráfego DNS. A principal contribuição desta metodologia é que para os principais componentes (leitura do tráfego, agregação em fluxos, extração dos grupos mais significativos e classificação das classes de comportamento) o funcionamento é único, ou seja, para uma mesma entrada, o resultado sempre será idêntico. No entanto, o processo de correlação e inspeção das classes de comportamento é dinâmico. Isso permite que novos componentes sejam incorporados na metodologia sem afetar o funcionamento principal. Portanto, outros pesquisadores podem contribuir com a metodologia proposta sem levar em consideração o funcionamento interno da metodologia.

Os resultados demonstram que a metodologia proposta é capaz de detectar um conjunto de anomalias que prejudicam o funcionamento da Internet. Além disso, a interpretação dos resultados confirma que, apesar da análise do tráfego DNS não ser uma atividade recente, seu estudo ainda é necessário, pois diversas anomalias de rede podem ser mitigadas através da análise do tráfego DNS. Um exemplo de anomalia são consultas do tipo MX por máquinas infectadas por aplicações de código malicioso.

Desta forma, este trabalho contribui de maneira significativa no processo de detecção de anomalias através do tráfego DNS, demonstrando que a combinação de informações do cabeçalho IP e das mensagens DNS, por meio da entropia, é possível identificar padrões de comportamentos maliciosos no tráfego.

1.5 Organização da dissertação

O restante deste trabalho está estruturado conforme descrito a seguir.

O Capitulo 2 descreve os conceitos fundamentais para o entendimento do protocolo DNS, a definição de alguns dos principais ataques de rede que utilizam do tráfego DNS e suas principais características. Além disso, este capítulo também apresenta os conceitos de teoria da informação usados neste trabalho.

O Capítulo 3 apresenta os trabalhos existentes e que foram utilizados como referência para o desenvolvimento da metodologia proposta.

O Capítulo 4 descreve a metodologia desenvolvida para detecção de atividades maliciosas na Internet, usando os conceitos apresentados no Capítulo 2. As principais etapas da metodologia – coleta do tráfego, agregação em fluxo de dados, extração dos grupos mais significativos e classificação de anomalias, são descritas em detalhes.

O Capítulo 5 apresenta algumas avaliações da metodologia proposta utilizando tráfego de rede da Internet fornecida pelo Centro de Operações OARC [9]. Os experimentos serão detalhados e também serão mostrados os resultados obtidos.

Por fim, o Capítulo 6 apresenta as conclusões e sugestões como trabalhos futuros.

2 Conceitos Básicos

"Embora ninguém possa voltar atrás e fazer um novo começo, qualquer um pode começar agora e fazer um novo fim."

Chico Xavier

O sistema de nomes de domínio (DNS) é utilizado por muitas aplicações como navegadores de Internet, aplicações de correio eletrônico e softwares de mensagens instantâneas. O serviço de tradução de nomes, na sua essência, associa nomes simbólicos a endereços numéricos IP, permitindo aos usuários utilizar recursos compartilhados em um ambiente de rede conectado.

Este Capítulo apresenta os principais fundamentos do sistema DNS e descreve alguns comportamentos maliciosos ou anomalias de rede que se utilizam do tráfego DNS como ponto de partida para comprometer ou interromper serviços de rede.

2.1 Sistema de Nomes de Domínio (DNS)

No projeto embrionário da Internet os usuários acessavam os serviços de rede a partir do endereçamento IP das máquinas, entretanto, essa abordagem apresenta desvantagens. Por exemplo, é mais fácil para seres humanos associar nomes simbólicos a objetos que a seqüências numéricas [10]. Essa limitação implica que a partir do crescimento do número de dispositivos conectados à rede, esse modelo de comunicação é inviável. Outro problema observado à época era a ausência de um controle distribuído que pudesse gerenciar a tabela de nós de rede TCP/IP [11]. Essa tabela armazena o endereço IP e o nome do nó em um arquivo texto que era copiado para outras redes.

O sistema de nomes de domínio, ou DNS, foi proposto para solucionar esses problemas. O DNS é um banco de dados distribuído, usado para traduzir nomes de máquinas para os seus respectivos endereços IP. O termo distribuído é aplicado na definição, pois nenhum nome de

domínio possui conhecimento sobre outros domínios existente na Internet, em outras palavras, cada domínio apenas responde pelo seu próprio espaço de nomes. Por exemplo, o domínio *google.com* não possui informações referentes ao domínio *yahoo.com* e vice-versa.

Utilizar uma base de dados distribuída permite que o acesso às informações não dependa de um nó central, essa característica garante maior escalabilidade do serviço de tradução de nomes. Além disso, o gerenciamento da base de dados é independente. Isso significa que o domínio *dcc.ufam.edu.br* pode ser administrado pelos alunos do curso de ciência da computação, enquanto que, o domínio *ufam.edu.br* seja administrado pelo quadro coorporativo da Instituição.

O DNS está formalizado nas RFCs 1034 e 1035 [12]. Sua função permite que as aplicações possam obter informações sobre os mais variados tipos serviços disponíveis na rede como a localização de servidores de correio eletrônico, servidores web, aplicações de comércio eletrônico (*e-commerce*), ou até mesmo outros servidores de nomes.

As seções seguintes apresentam o funcionamento do DNS e suas características que permitem compreender melhor a área de atuação deste trabalho.

2.2 Arquitetura do DNS

As informações do protocolo DNS são obtidas através de consultas ao próprio DNS buscando os dados sobre os domínios na arquitetura ilustrada na Figura 2.1. Esta hierarquia denota o nó principal como nó Raiz. Neste nível todas as consultas são encaminhadas ou direcionadas aos servidores de domínio de primeiro nível (TLD - *Top-level Domains*).

Domínio	**Funcionalidade**
.com	Destinado a domínios com fins comercias
.gov	Destinado a domínios do governo dos EUA
.us	Destinado a domínios dos EUA
.br	Destinado a domínios do Brasil

Tabela 2.1: Principais domínios e suas respectivas funções na árvore hierárquica do DNS.

Os domínios de primeiro nível foram definidos, inicialmente, por atribuições geográficas e funções organizacionais [11]. Cada domínio abaixo de um TLD é chamado domínio de segundo nível. Domínios atribuídos aos países são chamados de domínios de

primeiro nível de código de país ou ccTLD (*country-code-top-level domains*). A Tabela 2.1 sumariza uma lista de nomes de domínio de primeiro nível e sua descrição organizacional.

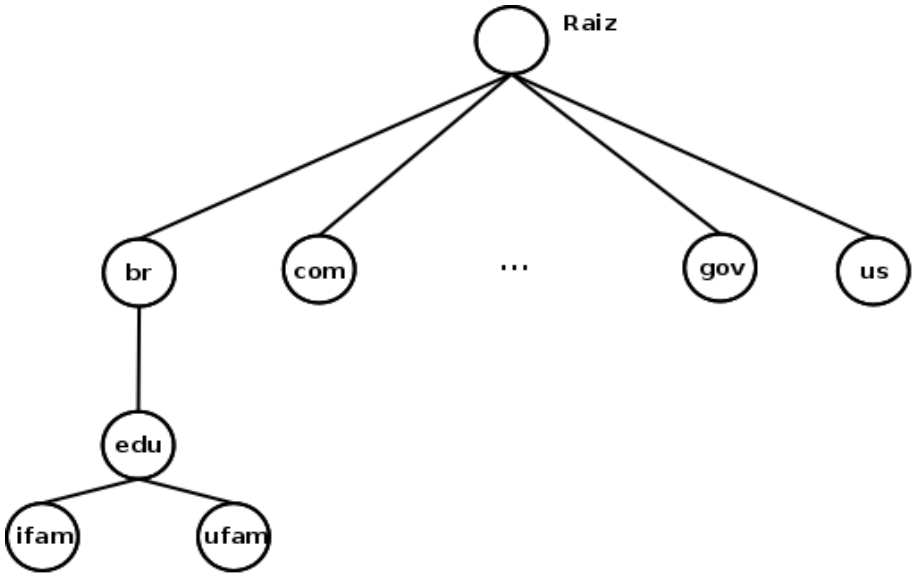


Figura 2.1: Exemplo da estrutura hierárquica do sistema de tradução de nomes.

Os nomes de domínios na árvore hierárquica do DNS denotam dois pontos fundamentais em sistemas baseado em nomes: a organização da estrutura de nomes e o espaço de nomes de domínio (*domain name space*) [11]. O espaço de nomes define a arquitetura dos nomes, como os nomes são registrados e consultados, isto é, como são construídos e interpretados no processo de resolução de nomes. O processo de resolução de nomes está descrito com mais detalhes na Seção 2.2.3.

2.2.1 Servidores de nome

O servidor de nome é uma aplicação que fornece o serviço de tradução de nomes em endereços, ou seja, mapeia os nomes de um domínio para endereços IP [11]. Além disso, o servidor de nomes é responsável por uma porção do espaço de nomes, por exemplo, *ufam.edu.br*. Esta porção também é conhecida como zona de domínio [13].

Para garantir disponibilidade dos dados, a estrutura do sistema de nomes de domínio permite que servidores de nome realizem atividades distintas. Existem vários tipos de servidores DNS, dentre os quais se destacam: servidores de nome raiz ou servidor raiz,

servidores de nomes de domínio de primeiro nível, servidores de nome com autoridade e servidores de nome recursivo.

Servidores de nome raiz são servidores com autoridade da zona raiz (ilustrada na seção anterior). Esses servidores apenas indicam ou referenciam aos clientes o caminho necessário que deve ser tomado para alcançar os servidores de nome com autoridade dos TLDs. Todo processo de resolução de nomes deve iniciar a partir de servidores de nome raiz, portanto, manter sua operação correta é crucial para o funcionamento da Internet. Assim, para garantir disponibilidade das informações existem 13 servidores lógicos que atendem todas as consultas na Internet [13]. É importante ressaltar que esses endereços utilizam *unicast* para reduzir tentativas maliciosas de ataques de rede

Os *domínios de primeiro nível* (TLDs) são atendidos pelos servidores de nomes de domínio de primeiro nível. Esses servidores são responsáveis pelos domínios *.gov*, *.com*, *.br* entre outros. É importante observar que nesse nível, assim como servidores raiz, as consultas são indicadas para alcançar os servidores de nome com autoridade. Em outras palavras, servidores de nomes de domínio de primeiro nível apenas informam o caminho para os servidores do nível abaixo.

Servidores de nomes com autoridade são servidores que possuem os endereços IP para os nomes de domínios sob sua jurisdição. Este tipo servidor pode responder por uma ou mais zonas de domínio. Por exemplo, considere a zona *ufam.edu.br*. Quando uma consulta é feita sobre essa zona, tanto o servidor primário ou secundário podem responder essa solicitação. A diferença entre esses servidores é que o primário obtém as informações da zona a partir da leitura de um arquivo localizado no seu próprio sistema de arquivo, enquanto que no servidor secundário, as informações são solicitadas do servidor primário [13].

*Servidor*es *de nome recursivo* apresentam uma função importante no processo de resolução de nomes, pois armazenam em *cache* as respostas obtidas dos servidores de nome com autoridade. Esse recurso reduz a sobrecarga de consultas que alcançam os servidores de nome raiz. Isso significa que, antes de consultar o servidor de nome raiz, o servidor de nome recursivo consulta sua base de dados pelo domínio solicitado. Caso o domínio não seja encontrado, o servidor de nome recursivo envia uma nova solicitação ao servidor de nome raiz [11].

2.2.2 Registros de recurso

Cada nome de domínio na árvore do DNS possui um atributo que permite identificá-lo no espaço de nomes de domínio. Os atributos são chamados registro de recurso, os quais fornecem um mapeamento entre os nomes de domínio e objetos de rede como, por exemplo, o endereçamento numérico IP, servidor de correio eletrônico e servidor de nome com autoridade. Cada mensagem de resposta DNS carrega um ou mais registros de recursos.

Existem diferentes registros de recursos que estão descritos com mais detalhes na RFC 1034. A Tabela 2.2 mostra alguns dos principais tipos registros de recursos.

Tipo	Significado	Exemplo
A	Indica um endereço numérico a partir do nome simbólico	dcc.ufam.edu.br, 200.17.49.2, A
PTR	Realiza a tradução reversa de nomes apontando para um nome simbólico por meio do endereço IP	dcc.ufam.edu.br, 2.49.17.200.in-addr.arpa, PTR
MX	Especifica servidor responsável pelo recebimento do correio eletrônico que chega ao domínio	dcc.ufam.edu.br, solimoes.dcc.ufam.edu.br, MX
NS	Especifica o servidor de nomes com autoridade pelo domínio	dcc.ufam.edu.br, uirapuru.ufam.edu.br, NS

Tabela 2.2: Tipos de registros de recursos

O formato do registro de recurso pode ser denotado como uma tupla contendo os seguintes campos: Nome, Tipo, Classe, Tempo de Vida, Tamanho dos Dados de Recurso e Dados de Recurso. O *Nome* define o nome de domínio. O campo *Tipo* indica o tipo de recurso do registro. A *Classe* apresenta a família do protocolo utilizada. O campo *Tempo de Vida* define o tempo que o registro deve ficar armazenado ou removido. O campo *Tamanho dos Dados de Recurso* especifica o tamanho de dados de recurso. Os *Dados de Recurso* apresentam o formato do registro de recurso, no entanto, essas características dependem da classe e do tipo de registro de recurso [11] [14].

2.2.3 Resolução de nomes

O processo de resolução de nomes é realizado quando um cliente DNS solicita ao servidor de nomes recursivo um endereço de domínio, ou o próprio cliente, através do resolvedor de nomes, realiza todo processo da consulta. Para ilustrar como o servidor de nome

recurso interage com a infraestrutura do DNS, a Figura 2.2 ilustra o processo de resolução de nomes através de um servidor DNS recursivo.

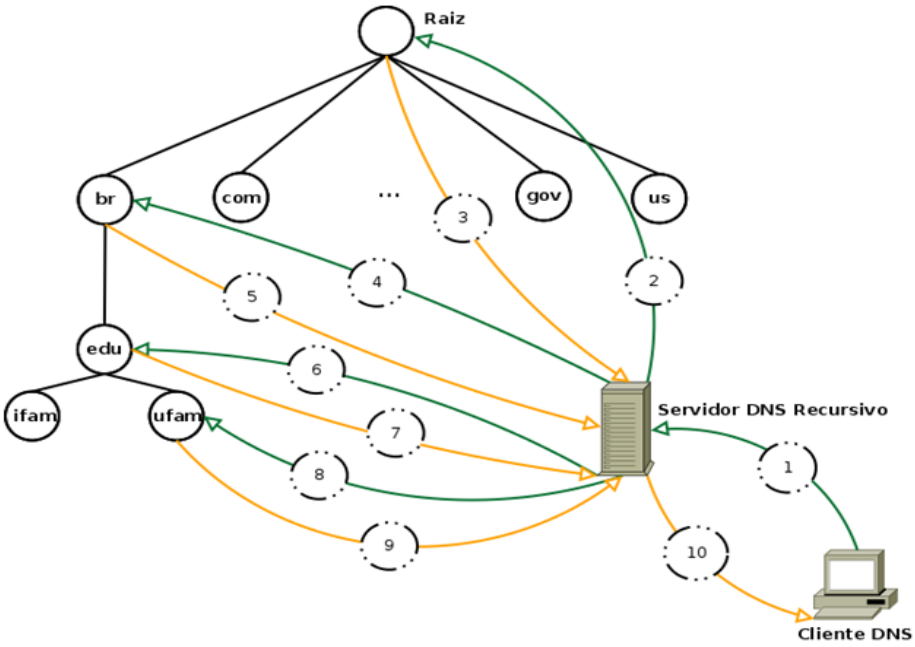


Figura 2.2: Processo de resolução de nomes através de um servidor de nomes recursivo.

No primeiro passo (1), o cliente DNS faz uma solicitação para um endereço de domínio, por exemplo, *www.ufam.edu.br*. No segundo passo, o servidor de nome recursivo recebe essa solicitação DNS e transmite a mensagem ao servidor de nome raiz (2). Como o servidor de nome raiz apenas possui conhecimento sobre os domínios de primeiro nível, o servidor de nomes recursivo recebe uma resposta indicando o servidor mais próximo que possui as informações sobre essa solicitação (3). Em seguida, o servidor de nome recursivo retransmite a solicitação aos servidores com autoridade pelo domínio *.br* (4) e recebe como resposta, uma referência dos servidores com autoridade do domínio *.edu.br* (5). Na etapa seguinte, o servidor recursivo solicita aos servidores de nome com autoridade do domínio *.edu.br* informações sobre o domínio *ufam.edu.br* (6), os quais respondem com o endereço do servidor de nome com autoridade sob o domínio *ufam.edu.br* (7). Finalmente, o endereço IP de *www.ufam.edu.br*, armazena esse dado no servidor de nome recursivo e encaminha a resposta para o cliente DNS para concluir o seu acesso (8) e (9).

2.2.4 Formato do Cabeçalho DNS

Toda comunicação do DNS, tanto consultas quanto respostas, são transmitidas em um único formato de mensagem. O cabeçalho da mensagem é encontrado nos primeiros 12 bytes do pacote, como ilustra a Figura 2.3.

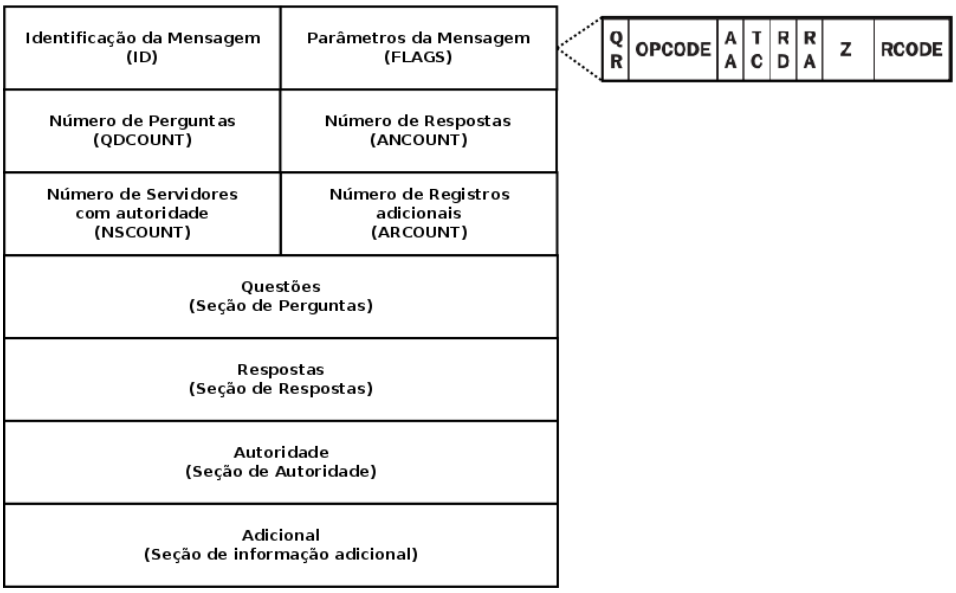


Figura 2.3: Formato do cabeçalho DNS.

Abaixo, segue uma breve descrição dos campos do cabeçalho do protocolo DNS.

- *Identificação*: o identificador da mensagem ou ID é responsável por identificar a consulta DNS. Esse campo é copiado nas respostas permitindo que o cliente reconheça a pergunta realizada.
- *Parâmetros da Mensagem*: esse campo utiliza bits para marcar opções na mensagem. Com um bit especifica se a mensagem é uma consulta (0) ou uma resposta (1). O código da operação (OPCODE) utiliza quatro bits para indicar o tipo da mensagem, como pergunta padrão (0), consulta inversa (1), pergunta de estado (2), pergunta de notificação (4) e pergunta de atualização (5). Quando uma mensagem é respondida pelo servidor com autoridade o campo AA é marcado com 1 bit. Um bit de recursão desejada (RD) é habilitado quando o cliente solicita ao servidor recursivo uma pergunta DNS. Um bit marcado na reposta identifica se a

recursão está disponível (RA). O código de resposta (RCODE) utiliza quatro bits para especificar se a resposta não teve erro (0) e aconteceu um erro no formato (1).

- *Número de perguntas*: o número de perguntas indica quantas perguntas está associada àquela pergunta.
- *Número de respostas*: esse campo registra a quantidade de respostas obtidas.
- *Número de servidores com autoridade*: o número de servidores com autoridade representa a quantidade de servidores que possuem autoridade sob o domínio consultado.
- *Número de informações adicionais*: esse campo indica quantos servidores adicionais podem ter conhecimento sobre o domínio consultado.
- *Questão*: esse campo contém informações sobre a solicitação, como o nome da consulta (QNAME), tipo da consulta (QTYPE) A, PTR, NS, e MX.
- *Resposta*: esse campo possui informações sobre os registros de recurso usados nas consultas e o tempo que dado será armazenado. No campo da resposta pode conter vários tipos de registros de recursos.
- *Autoridade*: esse campo armazena informações sobre os servidores de nome com autoridade.
- *Informações* Adicionais: esse campo contém informações sobre outros registros na resposta.

2.3 Vulnerabilidades do DNS

A formalização do protocolo DNS assume que as consultas DNS devem ser transmitidas em texto claro, na camada de rede. Em razão deste comportamento, o tráfego DNS pode ser interceptado e manipulado durante a comunicação entre o cliente e o servidor de nomes. Apesar do protocolo DNS apresentar essas características, esse padrão de comunicação não é considerado como vulnerabilidade do DNS [15].

Devido à importância do tráfego DNS para aplicações de rede, atacantes fazem uso das consultas como um dos principais pontos de partida para possíveis ataques de rede. Este trabalho aponta que as vulnerabilidades do DNS podem ser divididas em duas categorias:

ataques que exploram falhas de segurança no serviço de tradução de nomes e ataques de rede que utilizam o tráfego DNS como ponto de partida para outros ataques de rede. Na primeira situação, atacantes subvertem o funcionamento do protocolo para envenenar o resolvedor de um cliente ou servidor de nomes, a fim de direcioná-lo a um site comprometido. Na segunda categoria, atacantes utilizam o tráfego DNS para identificar possíveis alvos configurados em redes remotas, ou ainda, encontrar servidores de correio eletrônico para infectar outras máquinas conectadas à rede.

É importante mencionar que explorações no sistema operacional que hospeda o servidor de nomes devem ser consideradas, pois, tais ameaças podem ser mitigadas através de políticas de atualização de software.

2.3.1 Ataques que Exploram Vulnerabilidades do Protocolo DNS

Esta seção apresenta ataques que tentam subverter o funcionamento do DNS direcionando os clientes da rede para um site comprometido ou controlado por um atacante.

2.3.1.1 Interceptação de Pacotes

O protocolo DNS possui um comportamento bem definido, onde o cliente, através do resolvedor de nomes, emite solicitações ao servidor de nomes a fim de que este último responda suas requisições. No entanto, este processo de comunicação é vulnerável a técnicas de interceptação de pacotes onde o criminoso está localizado no mesmo segmento de rede entre o cliente e o servidor de nomes. Esse processo de interceptação permite que o atacante manipule as respostas com informações falsas para direcionar o cliente para sites comprometidos pelo atacante [16]. Esta abordagem de coleta de dados também é conhecida como *man-in-the-middle attacks* [17].

2.3.1.2 Ataques de Envenenamento de Cache

Ataques de envenenamento de cache são explorados a partir da predição da porta de origem UDP do cliente que originou a consulta e o número seqüencial que identifica a solicitação DNS (ID) [18].

A Figura 2.4 sumariza o conceito de um ataque de envenenamento de cache. O cliente (resolvedor *stub*) emite uma consulta ao servidor de nomes recursivo, que por sua vez, recebe

a resposta do servidor com autoridade pelo domínio consultado. Então, o servidor recursivo armazena a resposta no cache e retorna informações ao cliente. O envenenamento de cache é explorado pelo intruso enviando milhares de respostas forjadas para o servidor recursivo, antes que o servidor com autoridade envie sua resposta para o cliente que solicitou o endereço. O princípio deste abuso de rede está relacionado ao paradoxo do aniversário, onde duas ou mais pessoas em um grupo de 23 compartilham o mesmo dia do aniversário [19]. Assim, ao enviar uma quantidade considerada de respostas, é razoável afirmar que existe uma probabilidade P(A) de sucesso neste ataque.

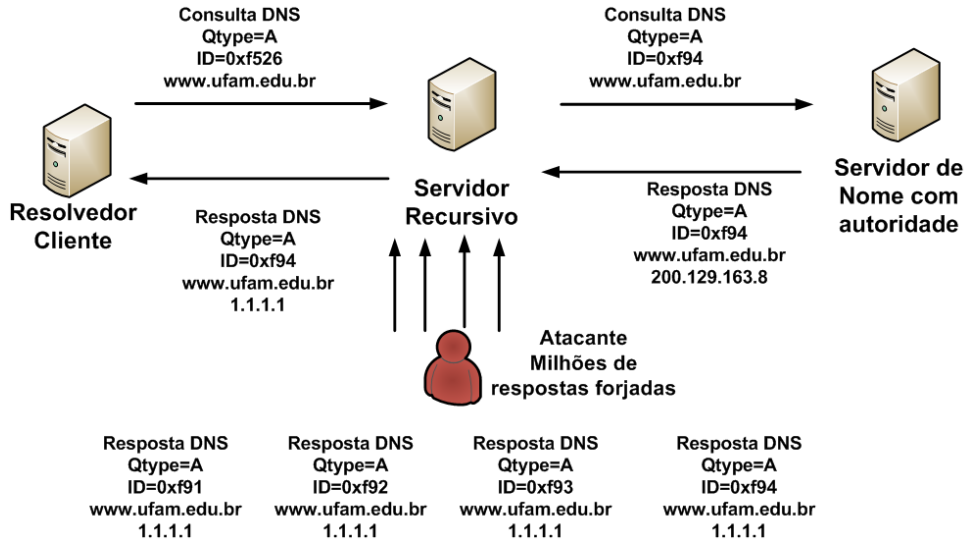


Figura 2.4: Esquema de um ataque de envenenamento de cache contra servidor de nomes recursivo.

Para tornar o entendimento mais claro, o padrão de comportamento apresentado na Figura 2.4 pode ser ilustrado através de uma representação gráfica usando *graphlet* [20]. Tal representação é útil, pois permite entender o padrão de comunicação na rede das possíveis anomalias. Por exemplo, considere a Figura 2.5. É possível observar um endereço IP de origem (*srcIP*) comunicando com um endereço IP de destino (*dstIP*) utilizando grande variação para porta de origem (*srcPort*), alta concentração para porta de destino (*dstPort*) e alta dispersão para o identificador da mensagem DNS. Por razões visuais, o tipo de registro de recurso e o nome da consulta são assumidos como fixos.

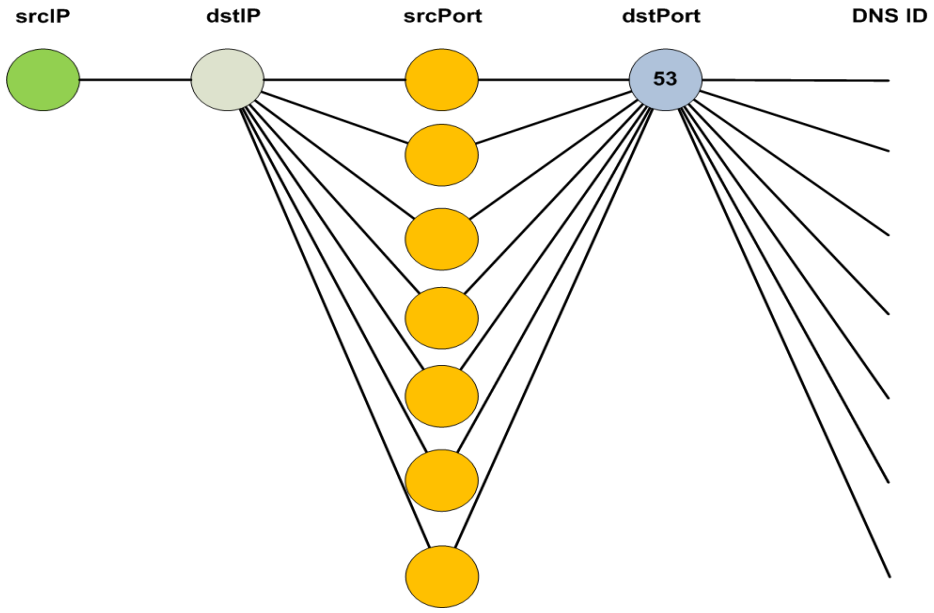


Figura 2.5: Esquema em *graphlet* de um ataque de envenenamento de cache.

2.3.1.3 Atualização dinâmica do DNS

Em redes de computadores, onde a atribuição de endereços IP é realizada automaticamente através do protocolo de configuração dinâmica de nó (*Dynamic Host Configuration Protocol*) [21], o cliente pode atualizar a zona de domínio com suas próprias informações (endereço IP, nome de domínio totalmente qualificado). No entanto, a possibilidade do cliente alterar a configuração da zona de domínio pode ser explorada por um atacante. Isso acontece porque essa modificação é baseada na confiança entre o cliente e o servidor [17]. Nesse contexto, é possível que o atacante utilize um conjunto de endereços IP falsos para interromper o serviço de tradução de nomes.

2.3.2 Ataques de rede através do tráfego DNS

Esta seção apresenta técnicas e atividades maliciosas que usam o tráfego DNS como ponto de partida para comprometer ou obter informações sobre possíveis alvos de ataques.

2.3.2.1 Ataques de negação de serviço

Ataques de negação de serviços podem ser utilizados para interromper o serviço de tradução de nomes, ou ainda, como um ponto de partida para cessar o fornecimento de uma aplicação. Por exemplo, em [22] é apresentado um ataque de negação de serviço que

amplifica o tamanho das respostas DNS. Nessa abordagem, o atacante formula solicitações para vários servidores de nomes recursivos na internet, cujas respostas, ultrapassam o limite de 1500 bytes da unidade de transmissão máxima (MTU) da rede. Essas perguntas possuem o endereço IP de origem forjado para que as respostas sejam redirecionadas para o alvo do ataque. Portanto, ao receber a quantidade de pacotes, o servidor aumenta a carga de processamento para atender todas as solicitações recebidas.

É importante ressaltar que servidores de nome raiz não atendem consultas recursivas, isto é, suas respostas apenas referenciam um servidor mais próximo que possui conhecimento sobre o domínio consultado. No entanto, é possível que o atacante direcione as respostas desse tipo de ataque para servidores de nome de raiz, interrompendo, desta forma, o serviço de tradução de nomes de domínio [23].

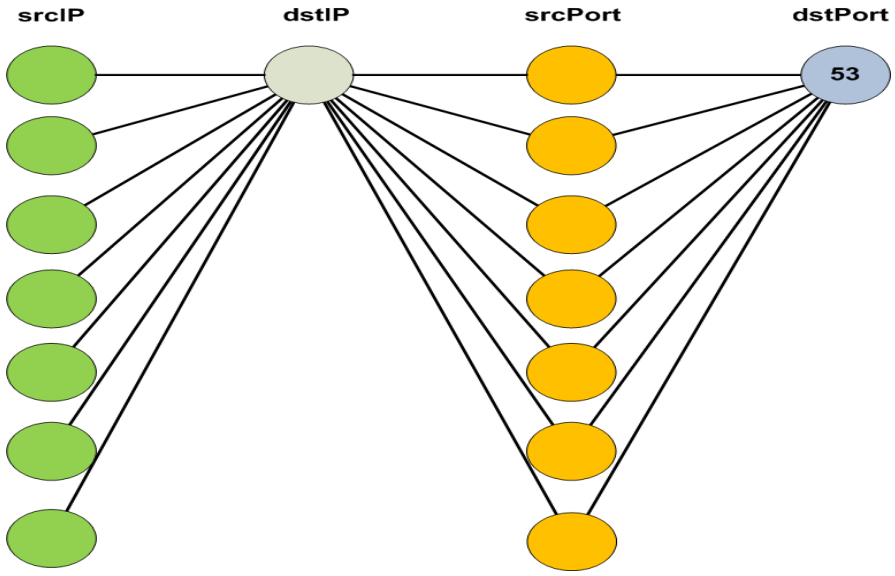


Figura 2.6: Esquema em *graphlet* de um ataque de negação de serviço (DoS)

Para tornar o entendimento mais claro, o padrão de um ataque de negação de serviço pode ser ilustrado através de uma representação gráfica usando *graphlet*. Apenas os componentes do tráfego mais relevante para ilustração são apresentados. É possível observar através da Figura 2.6 alta dispersão para endereços IPs de origem (*srcIP*) para um único endereço IP de destino (*dstIP*). Essa comunicação utiliza várias portas de origem (*srcPort*) para uma única porta de destino (*dstPort*).

Além disso, outros elementos do ataque devem ser observados. Por exemplo, a quantidade de respostas ou consultas recebidas em um intervalo de tempo. A relação entre número de respostas e a quantidade de consultas realizadas, também contribui no processo de detecção de anomalias. Em [24] é observado que ataques de amplificação de respostas DNS registram maior incidência de respostas recebidas em relação a quantidade de solicitações processadas.

2.3.2.2 Ataques de reconhecimento de rede através do DNS (*PTR-Scan*)

Ataques de reconhecimento consultam endereços de redes remotas através do tráfego DNS para identificar possíveis nós ativos ou configurados. Essa abordagem utiliza dois parâmetros: o registro de recurso do tipo PTR e o endereço IP de destino do alvo que será investigado. Nessa abordagem, o registro reverso é útil para o criminoso, pois o retorno dessas consultas fornece o nome de domínio totalmente qualificado (FQDN) a partir de um endereço IP [11]

Além disso, programas de código malicioso também empregam técnicas de reconhecimento de rede para identificar possíveis nós comprometidos. Essa estratégia tem como objetivo infestar outros computadores localizados na rede [25]. Dependendo da natureza do programa malicioso, o reconhecimento pode utilizar portas de serviço bem conhecidas ou infestar outros computadores com base na confiança entre eles.

IP de Origem e Porta de origem	IP de Destino e Porta de destino	Registro de Recurso	Nome da Consulta DNS
a.b.c.d.3200	200.160.0.10.53	PTR?	245.XXX.XXX.201.in-addr.arpa.
a.b.c.d.3200	200.160.0.10.53	PTR?	99.XXX.XXX.189.in-addr.arpa.
a.b.c.d.3200	200.160.0.10.53	PTR?	92.XXX.XXX.201.in-addr.arpa.
a.b.c.d.3200	200.160.0.10.53	PTR?	187.XXX.XXX.201.in-addr.arpa.
a.b.c.d.3200	200.160.0.10.53	PTR?	57.XXX.XXX.201.in-addr.arpa.

Tabela 2.3: Exemplo de um ataque de reconhecimento de rede utilizando o registro de recurso do tipo PTR para identificar nós ativos.

Para ilustrar como o tráfego DNS pode ser utilizado para reconhecer endereços de rede, considere o exemplo apresentado na Tabela 2.3. O endereço IP denotado como "*a.b.c.d*" consulta uma das instâncias com autoridade do domínio .br (a.dns.br - 200.160.0.10). Nesse

exemplo, a porta de origem está fixa (porta 3200). Contudo, em outras variações desse ataque, é possível observar maior aleatoriedade para esse campo. O tipo de registro de recurso do tipo PTR é o mais freqüente nessa abordagem, sendo a grande maioria nesse tipo de ataque. Os registros A e NS também são consultados, por exemplo, para localizar endereços de servidores com autoridade de domínios de Internet banda larga. Os nomes das consultas refletem endereços IPs de clientes de Internet banda larga ou clientes que ainda utilizam conexão discada para acessar a Internet.

De acordo com [26], o alto índice de registros do tipo PTR no tráfego de rede é decorrente de ataques de força bruta contra servidores que oferecem o serviço SSH ou de aplicações para filtrar mensagens não desejadas (anti-SPAM).

Para tornar o processo de varredura mais aleatório e, desta forma, reduzir um possível padrão de assinatura na rede, atacantes utilizam endereços pré-definidos de classe de redes que serão reconhecidas, sendo esses endereços escolhidos de forma. A Tabela 2.4 ilustra um exemplo de como as classes de endereços são definidas.

```
unsigned char classes[] = { 3, 4, 6, 8, 9, 11, 12, 13,
      14, 15, 16, 17, 18, 19, 20, 21, 22, 24, 25, 26,
      ...
      61, 62, 63, 64, 65, 66, 67, 68, 80, 81, 128,
      ....
      156, 157, 158, 159, 160, 161, 162, 163, 164,
      ...
      183, 184, 185, 186, 187, 188, 189, 190, 191,
      192, 193, 194, 195, 196, 198, 199, 200, 201,
      ...
      232, 233, 234, 235, 236, 237, 238, 239  };
```

Tabela 2.4: Exemplo de classes de endereços utilizadas por *worms* que empregam o tráfego DNS para reconhecer possíveis alvos de rede [25].

Para tornar o entendimento mais claro, as principais características de técnicas de reconhecimento de rede podem ser representadas a partir de *graphlet*. A Figura 2.7 ilustra um esquema de reconhecimento de rede através do tráfego DNS. Nessa representação é possível observar um endereço IP de origem (*srcIP*) consultando o servidor de nomes (*dstIP*) a partir de uma única, ou, poucas portas de origem (*srcPort*), para a porta de destino (*dstPortP*). Essas

consultas utilizam com maior freqüência o registro reverso. Esse registro busca diversos endereços IP no nome da consulta DNS (*Qname*). Os registros A e NS podem aparecer com observações únicas.

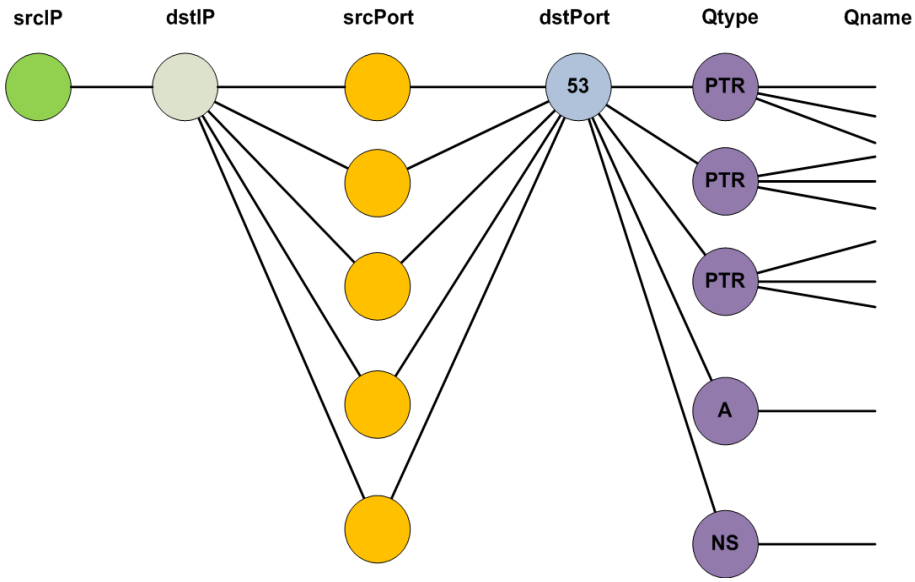


Figura 2.7: Esquema de um ataque de reconhecimento de rede através do registro de recurso do tipo PTR.

2.3.2.3 Propagação de mensagem em massa (SPAM)

Em [27] é apresentado que programas de código malicioso de propagação em massa por e-mail (*mass-mailing worms*) podem ser identificados a partir de consultas DNS do tipo MX. Em outras palavras, atividades maliciosas que utilizam esta abordagem, deixam assinaturas no tráfego de rede que facilitam sua detecção e mitigação.

Em contraste com o processo de resolução de nomes descrito na seção 2.2.3, um nó de rede comprometido por aplicações de código malicioso, solicita informações sobre servidores de correio eletrônico diretamente aos servidores de nome raiz. Desta forma, utilizando essa estratégia, é possível subverter sistemas de detecção de anomalias que monitoram consultas maliciosas em servidores recursivos [28] [29]. Essas anomalias são compostas por implementações que permitem realizar esse tipo de consulta, por isso, não dependem de um servidor de nome local para concluir a atividade. A Figura 2.5 ilustra esse processo.

A partir da observação dos padrões de comunicação entre os tipos de registro de recurso do tipo MX, PTR e A é possível detectar programas maliciosos de envio de mensagens em

massa [29]. Esses tipos de registros indicam fortes evidências de comportamento anômalo na rede. Por exemplo, em [30] uma variação do *worm Sobig* [31] é observada a partir da análise dos registros A e MX, sendo o registro que fornece o servidor de correio eletrônico o mais freqüente no tráfego. Portanto, é razoável afirmar que clientes consultando servidores de nome raiz apenas com os registros A e MX, sem consultar o registro reverso, podem ser considerados como infectado por algum programa de envio de mensagem em massa.

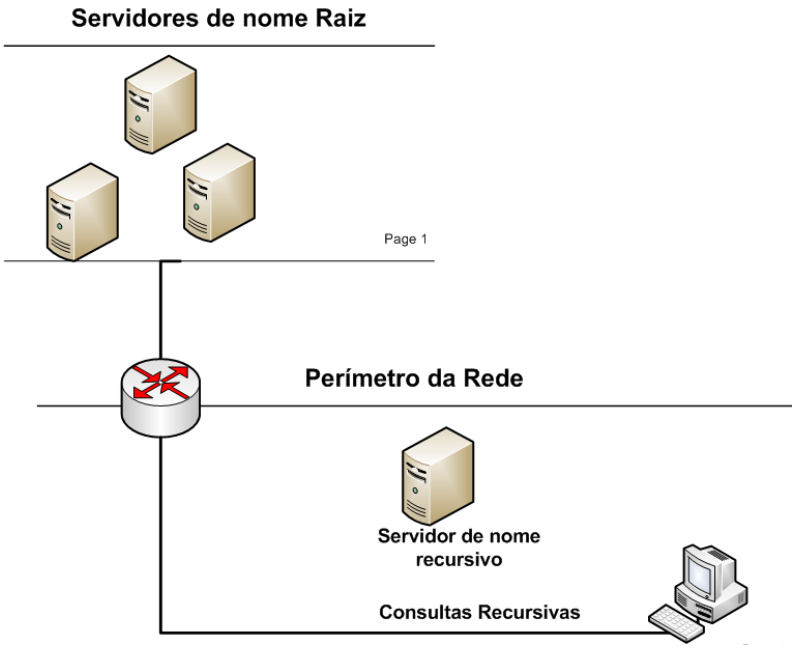


Figura 2.8: Exemplo de consulta de cliente DNS solicitando diretamente aos servidores de nome raiz.

A maioria de programas de propagação de mensagens em massa utiliza técnicas de engenharia social para enganar usuários a abrir suas mensagens eletrônicas (mails) ou executar arquivos em anexo. Essas abordagens sempre terão êxito, devido à falta de experiência de novos usuários utilizando sistemas de computador. Portanto, programas de propagação de mensagens em massa serão observados na rede com freqüência [32].

2.4 Teoria da informação – entropia

Essa seção descreve os principais conceitos sobre Teoria da Informação assim como sua aplicação no processo de detecção de anomalias de rede. Medidas da Teoria da Informação, mais especificamente, Entropia, tem sido utilizada com sucesso no processo de extração de

comportamentos que estão fora do padrão de comunicação. Esse fator positivo acontece porque o resultado da entropia fornece os comportamentos que mais se destacam no volume de tráfego analisado.

2.4.1 Entropia

O conceito de incerteza é muito mais amplo para ser definido, por isso, literaturas sobre Teoria da Informação [33] [34] ilustram exemplos que envolvem a incerteza em situações de tomadas de decisão. Por exemplo, quanto de incerteza existe em um lançamento de um dado não viciado? Quanto de escolha está envolvido na seleção de um evento, ou ainda, quanto de incerteza se tem conhecimento sobre um possível resultado [35]. No entanto, a quantidade de incerteza é reduzida quando uma informação pertinente é revelada. Por exemplo, considere o lançamento de um dado tendencioso que, em dez lançamentos, o lado seis teve sua face observada nove vezes. Em outras palavras, a incerteza nesta situação é praticamente nula.

Nesse contexto, a quantidade de incerteza está estritamente relacionada ao conceito de informação. Isso significa que, o montante de informações obtido pode ser mensurado a partir da quantidade de incerteza reduzida [33]. Portanto, para mensurar a quantidade de incerteza em uma informação, este trabalho utiliza os conceitos da Teoria da Informação.

Os fundamentos da Teoria da Informação, formalizados por Claude Shannon [35], têm sido estudados extensivamente devido sua aplicabilidade em diversas áreas com problemas teóricos, bem como em áreas específicas, tais como compressão de dados, transmissão e processamento de sinais, teoria dos ruídos, correção de erros, criptografia de dados e detecção de anomalias de rede. Dentre os princípios da teoria, a entropia apresenta uma importante característica para quantificação da informação contida numa mensagem.

A entropia é uma possível medida da informação, ou seja, na Teoria da Informação a entropia calcula a quantidade mínima de bits necessários para representar a fonte de uma mensagem. Portanto, a quantidade de informação contida numa mensagem é a quantidade de entropia necessária para representá-la [33].

A quantidade calculada pela entropia também pode ser interpretada como a medida da incerteza da mensagem. Por exemplo, considerando um espaço amostral de tempo finito e enumerável, onde eventos mutuamente exclusivos podem ocorrer, como "Hoje de manhã o

nível do Rio Negro subiu 0.1 centímetro" ou "Hoje de manhã o nível do Rio Negro subiu 0.2 centímetros". Esse tipo de dado, considerando o período das chuvas da região Amazônica, não fornece informação suficiente para deduzir algum comportamento. Contudo, o evento "Hoje de manhã o nível do Rio Negro subiu 10 metros" permite inferir mais detalhes desse comportamento. Por isso, muitos autores relacionam a informação como uma variável aleatória que dependendo da sua natureza, pode assumir N valores discretos [34] [35] [36].

Em teoria, a entropia é utilizada para medir quanto de incerteza uma variável aleatória pode assumir. Por exemplo, seja X uma variável aleatória discreta com o alfabeto χ e função massa de probabilidade definida $p(x) = \Pr\{X = x\}, x \in \chi$. Por conveniência, a função massa de probabilidade $p_X(x)$ é definida como $p(x)$. Portanto, a entropia $H(X)$ de uma variável aleatória é definida como [33]:

$$H(X) = -\sum_{x \in \chi} p(x) \log p(x) \tag{1}$$

Para tornar o entendimento mais claro, a Equação 1 pode ser ilustrada em um evento onde uma variável aleatória pode assumir 32 possíveis resultados uniformes. Para identificar um resultado, é preciso estabelecer um valor que represente os 32 valores. Então, a quantidade de bits necessária para representar esta informação equivale a 5 bits como ilustra a demonstração na Equação 2:

No contexto de análise de anomalias de redes, considere uma variável aleatória X que pode receber Nx valores discretos. Os valores Nx denotam possíveis quantias assumidas em análise (endereço IP de origem, endereço IP de destino, porta de origem ou tipo da consulta DNS). Além disso, é suposto que X seja observada por m vezes, a cada 5 minutos, por exemplo. Portanto, a probabilidade de $p(x_i)$ pode ser denotada por $p(x_i) = \frac{m_i}{m}, x_i \in X$, onde m_i é a freqüência ou o número de vezes que X recebe o valor x_i. Em outras palavras, $p(x_i)$ é a probabilidade de X assumir valores durante o intervalo de tempo medido.

$$
\begin{aligned}
H(X) &= -\sum_{i=1}^{32} p(i)\log p(i) \\
&= -\sum_{i=1}^{32} p\left(\frac{1}{32}\right) log\ p\left(\frac{1}{32}\right) \\
&= \log 32 \\
&= 5
\end{aligned} \tag{2}
$$

Os resultados que X pode receber está presente no intervalo entre $(0\log 0) = 0$ e $0 \leq H(X) \leq H_{max}(X) \coloneqq \log\{Nx, m\}$. Portanto, se a entropia mínima $H(X) = 0$, então o resultado indica concentração máxima, i.e., baixa entropia, pois a quantidade de bits para representar esta informação é reduzida. De forma análoga, em detecção de anomalias de rede, a entropia mínima indica que o endereço IP de origem é observado com mais freqüência no tráfego. Por outro lado, se a entropia máxima $H_{max}(X) = 2^{H_{max}(X)}$, então a distribuição dos eventos é equiprovável, isto é, existe dispersão máxima, todos os valores observados de X são diferentes ou únicos, com probabilidade $p(x_i) = \frac{1}{m}$. Considerando no caso de anomalias de rede, a entropia máxima indica que o endereço IP de origem está uniformemente distribuído.

2.4.2 Incerteza relativa

A entropia pode ser usada para avaliar o padrão de comportamento do tráfego de rede, caracterizando o comportamento do tráfego e determinando se a distribuição está concentrada ou dispersa. Portanto, para medir o grau de variação ou uniformidade da distribuição, a incerteza relativa (RU) faz uma generalização da entropia de Shannon, para indicar essas mudanças. O conceito de incerteza relativa aplicada à detecção de anomalias é incorporado do trabalho em [8].

Formalmente, a incerteza relativa é definida como:

$$
RU(X) \coloneqq \frac{H(X)}{H_{max}(X)} = \frac{H(X)}{\log\min\{N_x, m\}} \tag{3}
$$

Os resultados possíveis de RU(X) apresentam a medida do grau de aleatoriedade ou uniformidade dos valores observados em X. Em outras palavras, se $RU(X) = 0$ então existe

concentração máxima, isto é, todas as observações de X são do mesmo tipo, p(x)=1 para qualquer $x \in X$. De forma mais geral, deixe A denotar o (sub) conjunto de valores observados em X, i.e., $p(x_i) > 0$ para $x_i \in A$. Considere $m \leq N_x$. Então $RU(X) = 1$, se e somente se $|A| = m$ e $p(x_i) = \frac{1}{m}$ para cada $x_i \in A$. Em outras palavras, dispersão máxima, ou seja, todos os valores observados em X são diferentes ou únicos, o que permite inferir o índice mais alto de variação ou incerteza para os valores observados. Portanto, quando $m \leq N_x$, então $RU(X)$ fornece a medida aleatoriedade ou exclusividade dos valores observados em X.

Entretanto, quando $m > N_x$, então $RU(X) = 1$, se e somente se, $m_i = \frac{m}{N_x}$, logo $p(x_i) = \frac{1}{N_x}$ para $x_i \in A = X$, i.e., os valores observados estão uniformemente distribuídos sobre X. Neste caso, $RU(X)$ mede o grau de uniformidade dos valores observados de X. Como uma medida geral de uniformidade dos valores observados de X, a entropia condicional $H(X|A)$ e a incerteza relativa condicional $RU(\mathrm{X}|\mathrm{A})$ são condicionadas à X baseado em A. Então, a entropia $H(X|A) = H(X), H_{max}(X|A) = \log |A|$ e $RU(X|A) = H(X)/\log|A|$. Portanto, $RU(X|A) = 1$, se e somente se $p(x_i) = \frac{1}{|A|}$ para todo $x_i \in A$. Em geral, quando $RU(X|A) \approx 1$, os valores observados de X estão próximos a serem distribuídos uniformemente, assim, menos distinguíveis entre si. Enquanto que, $RU(X|A) \ll 1$ indica que a distribuição é mais distante da uniformidade, com poucos valores observados mais freqüentemente.

Em resumo, a incerteza relativa neste trabalho fornece o grau de variação ou uniformidade sem considerar o suporte ou tamanho da amostra. Esta medida de uniformidade é utilizada na Seção 4.3 para extrair os grupos mais significativos.

3 Trabalhos Relacionados

"A educação é a arma mais poderosa que você pode usar para mudar o mundo."

Nelson Mandela

O tráfego DNS é uma fonte importante para detecção de anomalias de rede, por isso, diferentes técnicas utilizam o tráfego DNS para identificar e classificar atividades maliciosas no tráfego. Tais técnicas, por exemplo, empregam algoritmos de mineração de dados para aprender novos padrões de comportamento. Outro exemplo observa as consultas do tipo MX para identificar máquinas infectadas por *mass-mailing worms*. Consultas com o domínio de primeiro nível inválido são detectadas a partir da análise passiva do tráfego. Modelos probabilísticos indicam características de ataques de negação de serviço distribuído no tráfego de rede.

Este seção descreve algumas dessas técnicas citadas e destaca os trabalhos relacionados que fazem uso do conceito de entropia para detectar anomalias de rede.

3.1 Detecção de anomalias através do tráfego DNS

O processo de detecção de anomalias através do tráfego DNS pode ser subdividido em duas categorias, poluição do tráfego e atividades maliciosas. A poluição do tráfego DNS é caracterizada por consultas que não deveriam alcançar os servidores de nomes, por exemplo, solicitações aos endereços definidos na RFC 1918 [37] ou consultas repetidas. Atividades maliciosas no tráfego DNS são geradas a partir de vírus ou ataques de rede.

Esta subseção descreve os principais trabalhos relacionados ao processo de detecção de poluição no tráfego DNS e atividades maliciosas no tráfego.

3.1.1 Poluição do tráfego DNS

O funcionamento da Internet depende da operação correta dos serviços de tradução de nomes. Como conseqüência, diversos trabalhos têm sido propostos para caracterização do

tráfego DNS na Internet [38] [39] [40]. Em [38] é apresentado um dos trabalhos pioneiros no processo de detecção de anomalias através do tráfego DNS. Nesse trabalho são avaliadas as características do DNS como cache de consultas e eficiência do resolvedor de nomes. Dentre essas observações, a quantidade de consultas repetidas e servidores de nome com problemas de configuração são indicados como os principais problemas de consumo indevido de recursos de rede. Em [41] a análise passiva do tráfego DNS dos servidores de nome raiz demonstra que anomalias observadas em [38], ainda são recorrentes. Além disso, esse trabalho apresenta novos comportamentos que não deveriam alcançar os servidores raiz como, por exemplo, consultas destinadas ao espaço de endereçamento da RFC 1918, solicitações com o bit de recursividade ativo, domínios de primeiro nível inválidos, ataques de negação de serviço e problemas de codificação em clientes DNS.

Para mitigar o volume de consultas inválidas que alcançam os servidores de nome raiz, em [39] existe uma proposta para que administradores de rede configurem uma zona de domínio local pare responder às consultas destinadas à zona de domínio *in-addr.arpa.* Outra solução encontrada em [40] sugere que fabricantes de software atualizem seus produtos com correções relacionadas ao DNS para reduzir a quantidade de consultas mal formadas.

Em [5] é demonstrado análise passiva do tráfego DNS coletado durante o projeto DITL [42] que visa, através de ações coletivas, monitorar o tráfego DNS de grandes servidores de nomes distribuídos ao redor do mundo. Os resultados demonstram que acima de 90% total do tráfego processado pelos servidores de raiz nos meses de janeiro de 2006, janeiro de 2007 e março de 2008, são consultas DNS que não deveriam ocorrer na Internet. Os autores observam ainda que consultas que empregam o registro de recurso do tipo A são mais freqüente no tráfego e representam 60% do total do tráfego. Este comportamento também é constatado em [43], que atribuem à grande parcela do registro do tipo A no tráfego às ferramentas de combate a mensagens não solicitadas (SPAM).

Recentemente, [44] apresenta uma análise passiva do tráfego DNS pertencente ao domínio .br coletado durante o projeto DITL [42]. Diferentemente, dos trabalhos em [41] e [5], os resultados demonstram que o registro do tipo PTR é o mais freqüente no tráfego, correspondendo a 42,91% do total de consultas. Enquanto que, a fração de consultas do tipo A representa 30,29% do total de tráfego analisado. Os autores atribuem esses resultados as atividades maliciosas como ataques de reconhecimento de rede através do registro PTR (*PTR-*

Scan), envio de mensagens não solicitadas em massa e o tráfego PTR gerado por soluções de combate a mensagens não solicitadas.

Além da poluição do tráfego DNS que consome os recursos da infraestrutura de Internet de maneira inapropriada, atividades maliciosas de rede também degradam e podem interromper os serviços de tradução de nomes. Por exemplo, em [45] é demonstrado como um ataque de negação de serviço (DoS) pode impactar nos serviços de tradução de nomes. Ataques de negação de serviço utilizam os servidores de nome raiz como refletores no ataque [41], onde várias consultas falsas são formuladas por atacantes para um alvo (usuário ou sistema).

3.1.2 Detecção de atividades maliciosas no tráfego DNS

Atividades maliciosas utilizam o tráfego DNS como ponto de partida para atacar outros computadores conectados à rede como, por exemplo, computadores infectados por vírus e vermes (*worms*) [27], envenenamento de cache [46], ataques de negação de serviço [47].

Para reduzir ataques de negação de serviço, a proposta em [3] monitora um grupo de atividades DNS que identificam redes controladas por atacantes (*bonets*). Esse monitoramento avalia padrões de comportamento dos *botmasters* no processo de agrupamentos dos nós de rede comprometidos (*bots*) pela aplicação maliciosa. Esse trabalho propõe um algoritmo que consegue distinguir entre atividades maliciosas, originadas por *botnets*, e comportamento de usuários válidos. O algoritmo considera que os nós comprometidos compartilham o mesmo padrão de comunicação como, por exemplo, número fixo de endereços IP que consultaram domínios de *botnets*. No entanto, esse trabalho utiliza um sistema gestor de base de dados (SGBD) para armazenar informações do tráfego de rede, desta forma, em redes de alta velocidade, essa abordagem é inviável.

Além de ataques de negação de serviço, redes controladas por atacantes são utilizadas para enviar SPAM em massa, comprometer outros nós conectados à rede ou coletar informações privilegiadas de usuário. Em [48], *worms* que tentam comprometer outros nós conectados à rede são detectados a partir da observação das respostas DNS. Esse trabalho considera que uma conexão de rede iniciada sem uma consulta DNS é uma atividade maliciosa. Essa premissa parte do princípio que usuários tendem a utilizar nomes simbólicos

que endereços IP. Entretanto, diferente deste trabalho, a proposta em [48] também observa as conexões TCP para inferir comportamentos maliciosos.

A redução de mensagem em massa não solicita tem sido proposta em diversos trabalhos [27][49] [50]. Em [27] é proposto à utilização de estimadores bayesianos no processo de mineração de dados. Esse trabalho possui conhecimento prévio de assinaturas de *mass-mailing worms*. Os resultados obtidos demonstram que o método proposto consegue reduzir em 89% o volume de mensagens não desejadas em um provedor de Internet do Japão. Por outro lado, em [49], não existe conhecimento prévio desse tipo de atividade. Nesse trabalho é proposta uma abordagem que utiliza séries temporais e máquinas de aprendizado não supervisionadas. A partir da representação de uma série temporal, é possível destacar duas categorias de usuário: usuários válidos e usuários infectados por *mass-mailing worms*. A inferência desse princípio é possível, pois máquinas infectadas por esse *worms* exibem comportamento similar na rede. No entanto, esse trabalho cria uma lista branca (*whitelist*) dos usuários legítimos que consultam freqüentemente os servidores de nome. Nesse contexto, devido ao volume de tráfego que deve ser analisado, essa abordagem é inviável quando o tráfego de um servidor de nome raiz ou servidor de TLD é observado. Em [50] *mass-mailing worms* são detectados a partir da observação dos fluxos TCP referente à porta de saída 25 dos servidores de correio eletrônico.

Outras abordagens sugerem que atividades maliciosas no tráfego podem ser detectadas a partir de assinatura de rede [51] [52]. Entretanto, essas soluções não são eficientes para mitigar anomalias ou ataques que são desconhecidos, em outras palavras, ataques tipo *zero-day* [53]. Para solucionar a dependência de regras de ataques, algumas abordagens empregam técnicas estatísticas para identificação, medição e caracterização de anomalias de rede. Por exemplo, em [54] o teorema de Chebyshev é utilizado para detecção de redes controladas por atacantes (*botnets*) através da análise do tráfego DNS. Esse cálculo permite inferir o desvio padrão em relação à média populacional da distribuição de dados. No entanto, essa abordagem não apresenta resultados satisfatórios quando comparado com métodos Bayesiano, que identificam redes controladas por atacantes com 95% de confiança nos resultados [55].

Em [56] máquinas de aprendizagem são utilizadas para identificar anomalias desconhecidas na rede através do tráfego DNS. Nesse trabalho o estimador Naïve Bayes é aplicado em uma base de dados para treinamento de detecção de anomalias. Os resultados

apresentam entre 60-95% de precisão de identificação de anomalias no tráfego. Entretanto, o estimador apresenta problemas durante a fase de classificação. Isto é, falso negativo, onde 22% de classes de ataques de rede são classificadas como tráfego WWW e 26% de tráfego web como ataque.

3.2 Teoria da Informação

Técnicas da teoria da informação, como a entropia, são utilizadas para medir a imprevisibilidade dos dados na rede [35]. No contexto de anomalias de rede, o uso da entropia permite agrupar um conjunto de informações, como endereço IP de origem (*srcIP*), endereço IP de destino (*dstIP*), porta de origem (*srcPrt*), porta de destino (*dstPrt*) e protocolo de rede, para criar perfis de padrões de comunicação de tráfego [8] [57]. Em [8], padrões encontrados revelam que 80% de varreduras com destino a portas TCP bem definidas como 135, 137, 138, são tentativas de exploração de vulnerabilidade por máquinas infectadas por algum tipo de vírus. A visualização do tráfego em fluxo detectou que 50% origem desse tráfego anômalo vem do continente europeu e asiático. O trabalho descrito em [8] serve como inspiração para realização deste trabalho, já que a metodologia proposta por Xu pode ser aplicada em outros contextos de detecção de anomalias de rede.

Em [58] é apresentado algumas técnicas da teoria da informação que podem ser empregadas na detecção de anomalias em tráfego de rede, tais como complexidade de Kolomogorov, entropia e entropia relativa. Nesse estudo, cada técnica é descrita conforme suas características. Por exemplo, os autores argumentam que o custo computacional da entropia básica é alto (exponencial) e que outras abordagens podem ser usadas como algoritmos de aproximação [59] e [60] para reduzir esta complexidade em tempo linear.

A análise de distribuição de recursos em [7] correlaciona informações bem conhecidas do cabeçalho IP como os endereços IP de origem e de destino, e portas de origem e de destino para identificar anomalias de rede. Os autores usam a entropia de Shannon como forma de avaliar o grau de concentração das distribuições de probabilidade dos números de portas e dos endereços IP.

Em [61] é aplicado análise de entropia com estimativa de Holt-Winters para detecção de anomalias no tráfego da Rede Nacional de Pesquisa (RNP). O método proposto nesse trabalho usa análise de entropia combinado com estimativas de Holt-Winters para potencializar o uso

das ferramentas tradicionais de gerência de rede para a detecção de certos tipos de anomalias, como os ataques de DDoS e os *worms*.

Outro exemplo do uso da entropia para detecção de eventos na rede é demonstrado em [62]. Nesse trabalho os recursos do tráfego de rede como endereço IP de origem (*srcIP*), endereço IP de destino, porta de origem, porta de destino e carga útil do pacote são correlacionados para identificar aplicações P2P (*peer-to-peer*) no tráfego de rede.

Outras abordagens de correlação de distribuição de recursos têm sugerido com sucesso a utilização do cabeçalho IP para detecção de anomalias em soluções baseadas em entropia. São exemplos desta técnica [53], [63] e [8]. Entretanto, a estratégia adotada neste trabalho diverge dos recursos escolhidos para detecção de anomalias através do tráfego DNS. Isso acontece porque as anomalias de rede que utilizam o tráfego DNS estão localizadas na carga útil da mensagem DNS.

4 Detecção de Anomalias Usando Entropia

"Primeiro eles te ignoram, depois riem de você, depois brigam, e então você vence."

Mahatma Gandhi

Este Capítulo apresenta a descrição da metodologia usada na detecção de anomalias de rede através do tráfego DNS. A metodologia proposta utiliza o conceito de entropia introduzido no Capítulo 2. O principal objetivo dessa metodologia é prover uma abordagem que permita aos operadores e pesquisadores compreenderem o comportamento e mudanças na dinâmica no tráfego de rede, sejam essas alterações, resultado de atividades intencionais ou não, como anomalias ou erros de configuração.

4.1 Metodologia proposta

A metodologia proposta está dividida em quatro etapas bem definidas. A primeira realiza a leitura do tráfego de rede e agrega os dados em fluxo de dados. Na segunda etapa são extraídos os grupos mais significativos baseados em padrões do tráfego coletado. Na terceira etapa, estes grupos são automaticamente classificados com base no seu comportamento. É importante destacar que essa classificação utiliza o cálculo da entropia relativa. A última etapa é responsável pela interpretação das classes de comportamento obtidas na etapa anterior. Para avaliar a precisão da metodologia proposta, técnicas de correlação dos dados são empregadas. A Figura 4.1 ilustra a organização da metodologia proposta.

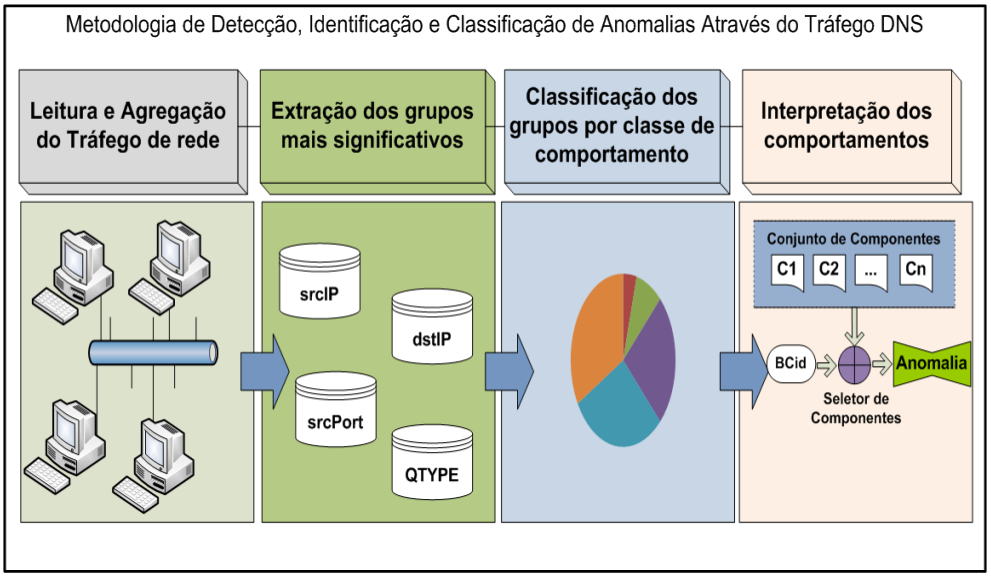


Figura 4.1: Etapas definidas para a metodologia proposta.

4.2 Leitura e agregação do tráfego de rede

O processo de leitura do tráfego pode ser realizado em tempo real (*online*) ou através captura de pacotes de rede, armazenados em um arquivo previamente salvo (*offline*). O procedimento de leitura do tráfego consiste em coleta passiva dos pacotes de rede, em outras palavras, não existe intervenção durante o processo de medição do tráfego.

O procedimento de captura de tráfego utiliza a biblioteca a *libpcap* [64], a qual oferece conjunto de rotinas de procedimentos que permitem capturar o tráfego de rede, por exemplo, acesso às interfaces nativas de rede do sistema operacional ou mesmo leitura de pacotes armazenados em arquivos pode ser obtido na biblioteca.

A leitura do tráfego de rede ocorre da seguinte forma. Cada pacote é lido de forma seqüencial através da função *pcap_next*, na mesma ordem em que o pacote é observado na rede, ou ainda, a partir de um arquivo no formato da libpcap. Caso a leitura do pacote não retorne algum erro de processamento, a função *pcap_next* indica um ponteiro para uma estrutura *pcap_pkthdr*, a qual armazena informações sobre o pacote capturado. Isso significa que o pacote é copiado do espaço do super usuário (*kernel*) para o espaço de usuário, permitindo assim, sua manipulação por outras rotinas [65].

A estrutura *pcap_pkthdr* é constituída da hora da captura do pacote, tamanho do pacote no enlace de rede e tamanho do pacote capturado. Estas informações são úteis no procedimento de agregação de fluxo de dados.

A Figura 4.2 ilustra o procedimento de leitura do tráfego de rede. Este procedimento recebe três parâmetros. O primeiro deles é a estrutura da própria implementação da biblioteca *libpcap*, o segundo contém um ponteiro para estrutura *pcap_pkthdr*, o terceiro fornece informações sobre o quadro capturado como, por exemplo, endereçamento IP, portas UDP e dados sobre a mensagem DNS. A partir das informações presente no pacote, o intervalo entre as linhas 1-4 define estruturas de dados do tipo *Ethernet*, IP, UDP e DNS respectivamente. Em seguida a hora inicial e final, tanto em microssegundos quanto em segundos, são extraídas nas linhas 5-8. O endereço IP de origem e destino, portas de origem e destino, nome e o tipo de registro de recurso da consulta DNS são obtidos no intervalo entre as linhas 9-14.

Leitura do tráfego de rede

```
AbreDadosPcap(Opções Pcap, Ponteiro de Estrutura pcap_pkthdr, Pacote)
1: Define Cabeçalho Ethernet;
2: Define Cabeçalho IP;
3: Define Cabeçalho UDP;
4: Define Cabeçalho DNS;
5: hora_inicial_ms <- Ponteiro de Estrutura pcap_pkthdr;
6: hora_final_ms <- Ponteiro de Estrutura pcap_pkthdr;
7: hora_inicial_seg <- Ponteiro de Estrutura pcap_pkthdr;
8: hora_final_seg <- Ponteiro de Estrutura pcap_pkthdr;
9: endereço_IP_origem <- ExtraiInformaçãoIP(Cabeçalho IP);
10: endereço_IP_destino <- ExtraiInformaçãoIP(Cabeçalho IP);
11: porta_de_origem <- ExtraiInformaçãoUDP(Cabeçalho UDP);
12: porta_de_destino <- ExtraiInformaçãoUDP(Cabeçalho UDP);
13: nome_da_consulta <- ExtraiCargaUtilDNS(Cabeçalho DNS);
14: registro_recurso <- ExtraiCargaUtilDNS(Cabeçalho DNS);
15: fluxo <- ProcuraFluxo(endereço_IP_origem, endereço_IP_destino, porta_de_origem,
registro_recurso);
16: se fluxo = NULO então
17:     fluxo <- CriaFluxo(endereço_IP_origem, endereço_IP_destino, porta_de_origem,
porta_de_destino,
 registro_recurso, hora_inicial_ms, hora_final_ms, hora_inicial_seg, hora_final_seg);
```

18: InsereFluxoTabelaEspalhamento(fluxo);

19: se fluxo != NULO então

20: AtualizaQuantidadePacotes(fluxo);

21: AtualizaQuantidadeBytes(fluxo);

Figura 4.2: Procedimento de captura de tráfego de rede

Paralelamente ao procedimento de leitura de pacotes da rede, ocorre o procedimento de geração de fluxos de dados. Este procedimento, no contexto de medição de tráfego de rede envolve um conjunto de pacotes que passam por um ponto de observação na rede durante um determinado intervalo de tempo, compartilhando propriedades em comum [66]. Em outras palavras, a definição do fluxo é denotada como uma seqüência unidirecional de pacotes com algumas características em comuns que passam por um dispositivo de rede como endereços IPs, quantidade de pacotes e bytes, portas de origem e destino e o momento que o pacote foi observado na rede [67].

Na metodologia proposta, o fluxo de dados é composto por informação do cabeçalho IP (endereço de origem e destino), da camada de transporte UDP (porta de origem) e a carga útil do protocolo DNS (tipo de registro de recurso). A Tabela 4.1 sintetiza os componentes do fluxo de dados escolhidos e como serão referenciados durante a realização deste trabalho.

Camada de Rede (IP)		**Camada de Transporte (UDP)**	**Camada de Aplicação (DNS)**
Endereço de origem (srcIP)	Endereço de destino (dstIP)	Porta de Origem (srcPort)	Tipo do registro de recurso (QTYPE)

Tabela 4.1: Componentes do fluxo de dados.

As informações da camada de rede permitem entender o comportamento de comunicação entre os nós na rede, ou seja, cliente e servidor DNS. O tipo de registro de recurso permite identificar comportamentos maliciosos que utilizam o tráfego DNS como, por exemplo, ataques de reconhecimento de rede, solicitações mal formadas ou envio de mensagens não solicitadas.

Em comparação com a proposta original em [8], é importante ressaltar que as anomalias de rede que utilizam o tráfego DNS geralmente estão localizadas na carga útil do pacote DNS. Isso significa que, extrair informações como o tipo de registro de recurso, agrega mais valor

ao processo de correlação dos componentes do fluxo de dados que a porta de destino. Essa conclusão pode ser constatada através do conceito de entropia, isto é, como as consultas DNS são destinadas a uma única porta, 53, a entropia necessária para representar essa informação é baixa, já que não existe incerteza para esse dado.

Os campos descritos na Tabela 4.1 podem ser representados como uma estrutura de dados. A Figura 4.3 ilustra tal estrutura.

Fluxo de dados (em linguagem C)

```
// Estrutura do Fluxo de Dados
struct _flow {
        unsigned int        identifier;         /*!< Identificador do fluxo */
        struct in_addr          src_ip;     /*!< endereço IP de origem */
        struct in_addr          dst_ip;     /*!< endereço IP de destino */
        unsigned short          source_port;        /*!< porta de origem */
        unsigned char           ip_protocol;        /*!< TCP/UDP/ICMP */
        unsigned long           packet_count;       /*!< Total de pacotes no fluxo */
        unsigned long           bytes_count;        /*!< Total de bytes no fluxo */
        time_t                  ini_sec;            /*!< Tempo inicial – segundos */
        time_t                  ini_mic;            /*!< Tempo inicial - microssegundos */
        time_t                  end_sec;            /*!< Tempo final – segundos */
        time_t                  end_mic;            /*!< Tempo final - microssegundos */
        unsigned short          qtype;              /*!< Registro de recurso */
};
typedef struct _flow Flow_t;
```

Figura 4.3: Estrutura de dados de um fluxo de dados.

Cada fluxo de dados gerado é armazenado em uma estrutura de dados eficiente, denominada de tabela de dispersão. No contexto de medição de tráfego de rede, as tabelas de dispersão são utilizadas com sucesso devido ao tempo de acesso rápido às informações desejadas [68] [69]. No pior caso, tabelas de dispersão que usam listas encadeadas para resolver os problemas de colisão, o tempo é $\Theta(n)$, por outro lado, se a função de *hash* dispersar as chaves entre os espaçamentos de maneira eficiente, o tempo é $O(1)$ [70].

Por último, após o processo de leitura do tráfego de rede e geração dos fluxos de dados, é necessário gerar um arquivo texto com os fluxos inseridos na tabela de dispersão. Esse

procedimento é necessário, pois os fluxos de dados estão alocados em memória RAM. Cada linha desse arquivo representa um fluxo de dados processado. O arquivo contendo todos os fluxos de dados armazenados é repassado para a segunda etapa da metodologia, que por sua vez, aplica um algoritmo de extração de grupos mais significativos observados no tráfego analisado.

4.3 Extração dos grupos mais significativos

O método proposto de detecção de anomalias através do tráfego DNS utiliza os tipos de registro de recurso que estão localizados na carga útil da mensagem. A combinação de informações do cabeçalho do pacote e da carga útil possibilita a compreensão mais detalhada da dinâmica de comunicação na rede.

Algumas abordagens recentes têm classificado anomalias de rede através da análise de fluxos de dados [8] [7]. O processo de detecção envolve o conceito de entropia para extrair mudanças no tráfego com base nos componentes do fluxo de dados. Estes componentes também são considerados como recursos do tráfego.

A metodologia proposta empresta os conceitos de extração de grupos significativos proposto por Xu [8] e correlaciona os componentes do tráfego DNS com recursos do cabeçalho IP para detectar anomalias de rede através do algoritmo de aproximação descrito na Seção 4.5. As subseções seguintes apresentam como os recursos do tráfego são correlacionados na geração de grupos significativos.

4.4 Definição das dimensões de cada grupo chave

A definição do grupo chave envolve a observação de uma característica importante do fluxo de dados em relação a outros recursos do tráfego. Por exemplo, é possível monitorar o comportamento de um nó em relação ao seu padrão de comunicação na rede como porta de destino, endereços IP ou tipo de registro de recurso. Outro exemplo dessa correlação observa a variação de conexões (endereços IP distintos) recebidas por determinada porta de serviço como, por exemplo, DNS. Em outras palavras, um grupo chave possui um recurso do tráfego fixo que é correlacionado com outras características do fluxo de dados. Essa correlação dimensional permite inferir comportamentos sobre tráfego analisado. A Figura 4.4 ilustra como um grupo chave pode ser correlacionado com cada dimensão de recurso do tráfego.

O processo de correlação dos recursos do tráfego ocorre em um espaço quadridimensional, onde cada grupo chave é confrontado com três dimensões distintas, X, Y, Z, respectivamente. Esta relação resulta em quatros grupos chaves de comportamentos.

É importante enfatizar que a proposta original em [8] correlaciona somente informações que estão localizadas no cabeçalho IP e na camada de transporte. No entanto, a metodologia proposta neste trabalho utiliza informações da camada de aplicação, especificamente do protocolo DNS, para inferir comportamentos anômalos na rede.

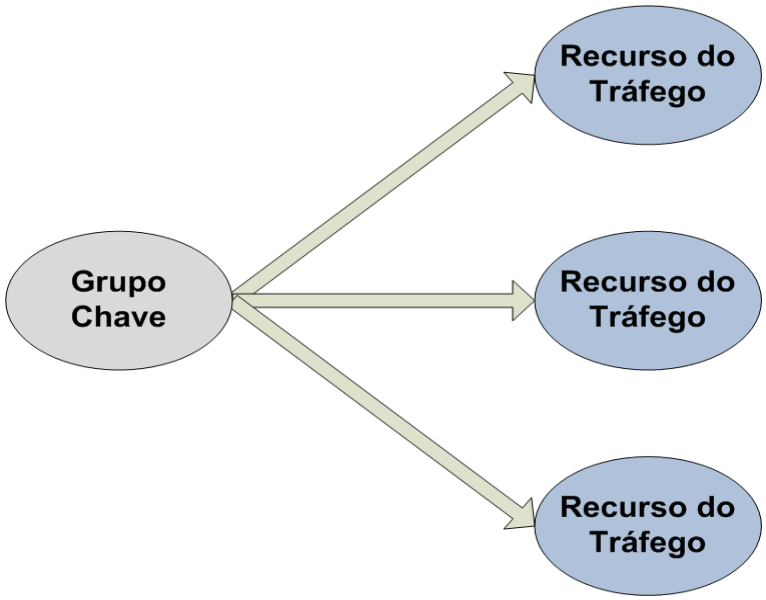


Figura 4.4: Exemplo como um grupo chave se relaciona com uma dimensão de recurso de tráfego de rede.

A Tabela 4.2 ilustra o grupo chave em relação às dimensões escolhidas. As colunas representam os recursos de tráfego (dimensões) e as linhas representam as observações de cada grupo chave. Para tornar o entendimento mais claro, o grupo chave endereço IP de origem (*srcIP*) é correlacionado com as dimensões porta de origem (*srcPort*), tipo do registro de recurso (QTYPE) e endereço IP de destino (*dstIP*). Uma possível interpretação desse comportamento seria a observação de como os tipos de registros de recursos são utilizados por esse endereço IP de origem. Algumas anomalias, por exemplo, reconhecem endereços IP válidos em uma rede através do tipo de recurso PTR (*PTR-scan*).

A correlação de recursos do tráfego é uma abordagem interessante para detecção de anomalias de rede, pois a grande maioria das anomalias compartilha a mesma estrutura de

recurso de tráfego de rede, isto é, mesmo usando algoritmos simples de agrupamentos, anomalias são detectadas em grupos distintos [7]. Além disso, um dos pontos positivos desta metodologia proposta é que através dos recursos do DNS, novos comportamentos maliciosos podem ser detectados a partir de mudanças no aspecto de distribuição nos recursos do tráfego.

Grupo Chave	**Dimensões**		
	X	**Y**	**Z**
srcIP	srcPort	QTYPE	dstIP
dstIP	srcPort	QTYPE	srcIP
srcPort	QTYPE	srcIP	dstIP
QTYPE	srcPort	srcIP	dstIP

Tabela 4.2: Relação do grupo chave com suas respectivas dimensões X, Y e Z.

Uma vez formalizado a definição do grupo chave e suas respectivas dimensões de correlação, é preciso destacar apenas os grupos mais significativos para detecção de anomalias. Em outras palavras, os grupos chaves que são distintos em termos de distribuição de probabilidade são considerados como significativos e extraídos da distribuição, sendo que este processo é repetido até que os grupos restantes sejam indistinguíveis um de outro, conforme [8]. O algoritmo de extração de grupos é descrito a seguir.

4.5 Algoritmo de aproximação

A extração dos grupos mais significativos considera cada grupo chave, *srcIP*, *dstIP*, *srcPort* ou *QTYPE* em relação as suas dimensões para extrair os elementos mais significativos. Por exemplo, os grupos *srcIP* e *dstIP* extraídos indicam um conjunto relevante de padrões de comportamentos dos nós da rede (padrões de comunicação). Enquanto que, a porta de origem e o tipo de registro de recurso estão relacionados ao comportamento do tráfego DNS. Esta subseção apresenta como a medida de incerteza relativa (condicional) pode ser utilizada para extrair os grupos mais significativos.

Considere uma dimensão X, por exemplo, *srcIP*, e um intervalo de tempo T, sendo m o total número de fluxos observados durante o intervalo de tempo, e $A = \{ a_1, \dots, a_n \}$, $n > 2$, como o conjunto de valores distintos de X (endereços IPs de origens, neste caso) que os fluxos observados podem assumir. Então a distribuição de probabilidade (induzida) P_A em X é dada

por $p_i \coloneqq P_A(a_i) = \frac{m_i}{m}$, onde m_i é o número de fluxos que assumem o valor de a_i, ou seja, recebendo o *srcIP* de a_i. Então a incerteza relativa (condicional), $RU(P_A) \coloneqq RU(X|A)$, mede o grau de uniformidade das características observadas em A. Seja β um valor próximo a 1, por exemplo, se $RU(P_A) \geq \beta$, então é possível deduzir que os valores observados estão próximos de serem uniformemente distribuídos, praticamente indistinguíveis. Por outro lado, existem valores prováveis em A que estão distantes dos padrões encontrados quando comparados com o restante. Portanto, é possível assumir que um subconjunto S de A é composto dos valores mais significativos de A, se S é o menor subconjunto de A considerando as seguintes premissas: i) a probabilidade de qualquer valor de S é maior que os valores restantes; e ii) a distribuição de probabilidade (condicional) do conjunto de valores restantes, isto é, $R: = A - S$, está próxima de ser uniformemente distribuída, ou seja, $RU(P_R) \coloneqq RU(X|R) > \beta$. Logo, S é composto dos valores com características mais significativas de A, enquanto que, os valores remanescentes são praticamente indistinguíveis uns dos outros.

Para ilustrar o conjunto de elementos em S, ordenados pelas características de valores de A, com base nas probabilidades assumidas, considere $\hat{a}_1, \hat{a}_2, \dots, \hat{a}_n$ tal como $P_A(\hat{a}_1) \geq P_A(\hat{a}_2) \geq \cdots P_A(\hat{a}_n)$. Então $S = \{\hat{a}_1, \hat{a}_2, \dots, \hat{a}_{k+1}\}$, e $R = A - S = \{\hat{a}_k, \hat{a}_{k+1}, \dots, \hat{a}_n\}$ onde k é o menor inteiro tal que $RU(P_R) > \beta$. Seja $\alpha^* = \hat{a}_{k+1}$. Então α^*equivale ao limite máximo de corte tal que a distribuição de probabilidade (condicional), no conjunto de valores restantes de R, esteja próxima de uma distribuição uniforme. Portanto, para extrair S de A leva-se em consideração que poucos valores (respeitando n) possuem probabilidades grandes, em outras palavras, o tamanho de S é pequeno, enquanto que, os valores restantes estão próximos de uma distribuição uniforme. Sendo assim, um limite ótimo de corte α^* pode ser procurado de maneira eficiente.

A Figura 4.5 apresenta o algoritmo de aproximação para extração de grupos significativos em S de A. Este trabalho considera a sugestão em [8] de $\alpha_0 = 2\%$ como valor inicial do algoritmo. Em seguida, o limite ótimo de corte α^* é procurado através de uma aproximação exponencial, onde o limite α é reduzido decrementando o fator de potenciação $\frac{1}{2^k}$ na *k*-ézima iteração. Enquanto a incerteza relativa da distribuição de probabilidade (condicional) P_R dos valores (restantes) do conjunto S for menor que β, o algoritmo examina cada valor em S e inclui aqueles cujas probabilidades excedem limite α no conjunto S dos

valores mais significativos. O algoritmo finaliza quando a distribuição de probabilidade dos valores restantes estiver próxima de uma distribuição uniforme, em outras palavras, quando a distribuição estive maior que β. Portanto, o último limite de corte encontrado pelo algoritmo é denotado como $\hat{a}^*$.

Extração dos grupos mais significativos baseados em Entropia

1: Parâmetros: $\alpha := \alpha_0$; $\beta := 0.9$; $S := \emptyset$;
2: Inicialização: $S := \emptyset$; $R := A$;
3: Calcula a distribuição de probabilidade PR e o RU $\theta := RU(PR)$;
4: Enquanto $\theta \leq \beta$ *faça*
5: $\alpha = \alpha \times 2^{-1}$;
6: para cada $a_i \in R$ *faça*
7: se $PA(a_i) \geq \alpha$ *então*
8: $S := S \cup \{a_i\}$; $R := R - \{a_i\}$;
9: fim se
10: fim para
11: calcule dist. prob. (cond.) PR e $\theta := RU(PR)$
12: fim enquanto

Figura 4.5: Algoritmo de aproximação para extração dos grupos significativos [8].

4.6 Classificação dos grupos em classe de comportamento

Após a finalização do processo de extração de grupos mais significativos, é necessário classificar os grupos com base nas dimensões correlacionadas. Este processo envolve o cálculo da incerteza relativa, descrita na Seção 2.4.2. Para ilustrar como esse conceito pode ser aplicado à medição e detecção de anomalias, considere o grupo chave *srcIP* sendo correlacionado com as dimensões porta de origem (*srcPort*), tipo da consulta DNS (QTYPE) e endereço IP de destino (*dstIP*). Essas dimensões são referidas como X, Y e Z respectivamente. Cada dimensão correlacionada pode receber qualquer valor, isto é, valores assumidos pela distribuição de probabilidade. Este conjunto de dimensões é definido em um vetor de incerteza relativa RU $[RU_x, RU_y, RU_z]$ como proposto em [57].

Uma alternativa para agregar os grupos mais significativos que apresentam comportamentos similares é por meio de associação de níveis a cada dimensão RU. As associações RU estão subdividas em três níveis discretos: 0 (baixo), 1 (médio) e 2 (alto) conforme ilustra a Equação 4.

$$\boldsymbol{L(RU)} = \begin{Bmatrix} \boldsymbol{0(baixo)}, & \boldsymbol{se\ 0 \leq RU \leq \varepsilon} \\ \boldsymbol{1(médio)}, & \boldsymbol{se\ \varepsilon \leq RU \leq 1-\varepsilon} \\ \boldsymbol{2(alto)}, & \boldsymbol{se\ 1-\varepsilon \leq RU \leq 1} \end{Bmatrix} \quad (4)$$

Para os endereços IP de origem e endereço IP de destino, *srcIP* e *dstIP* respectivamente, o valor estabelecido para $\varepsilon = 0.3$, enquanto para porta de origem (*srcPort)* e tipo do registro de recurso DNS (QTYPE) $\varepsilon = 0.2$. Os valores assumidos para ε incorporam as observações em [57]. O processo de associação permite criar 27 classes de comportamentos (BC – *behavior class*) distintas, isto é, o vetor $\left[L(RU_{x)}, L(RU_y), L(RU_z)\right] \in \{1,2,3\}^3$.

É importante considerar o valor do estimador quanto à porta de origem e tipo de registro de recurso, pois o atributo fornecido em [8] observa o comportamento de portas e aplicações, no entanto, este trabalho observa o comportamento das consultas emitidas pelos os nós na rede.

O vetor estabelecido $\left[L(RU_{x)}, L(RU_y), L(RU_z)\right]$ pode ser relacionado como um identificador inteiro denotado pela Equação 5.

$$BC_{id} = L(RU_X) * 3^2 + L(RU_Y) * 3^1 + L(RU_Z) * 3^0 \in \{0,1,\dots,26\} \quad (5)$$

Neste contexto, considere o grupo chave *srcIP* adotando o valor para classe de comportamento $BC_{id} = 3$, isto é, $BC_{id} = 0 * 3^2 + 1 * 3^1 + 0 * 3^0$. Este resultado pode ser traduzido da máxima concentração para a porta de origem, média variação para o tipo de registro de recurso e alta concentração para o endereço IP de destino, ou seja, $BC_{id=3} = [0,1,0]$. Outro exemplo ilustrado apresenta o comportamento do grupo chave tipo de registro de recurso (QTYPE). Considere a classe de comportamento $BC_{id} = 21$, esse resultado demonstra máxima dispersão para porta de origem, média concentração para endereço IP de origem e máxima concentração para endereço IP de destino, isto é, $BC_{id=21} = [2,1,0]$.

Para interpretar o resultado de uma classe de comportamento, é necessário correlacionar características do fluxo de dados. Essa correlação, com base no resultado das classes de comportamentos, permite inferir anomalias de rede. O processo de correlação é descrito na seção a seguir.

4.7 Interpretação das classes de comportamento

A última etapa da metodologia proposta envolve interpretação dos comportamentos resultantes da fase anterior. Em outras palavras, essa fase interpreta as classes de comportamento indicadas no processo de classificação dos grupos. Por exemplo, dado uma classe de comportamento, métricas ou componentes são utilizadas para analisar as características da classe. Uma vez encontrado comportamentos anômalos a partir dos padrões de comunicação, é possível inferir a natureza da anomalia na rede.

A Figura 4.6 ilustra como o processo de correlação e interpretação das classes de comportamento ocorre. As classes de comportamento, obtidas da fase anterior, são repassadas para um seletor que aciona um ou mais componentes que auxiliarão na definição do tipo de anomalia de rede. A seleção de componentes é baseada em regras estáticas, do tipo "*if-then-else"*, que foram estabelecidas a partir da análise passiva do tráfego DNS, a ser apresentada na Seção 5.

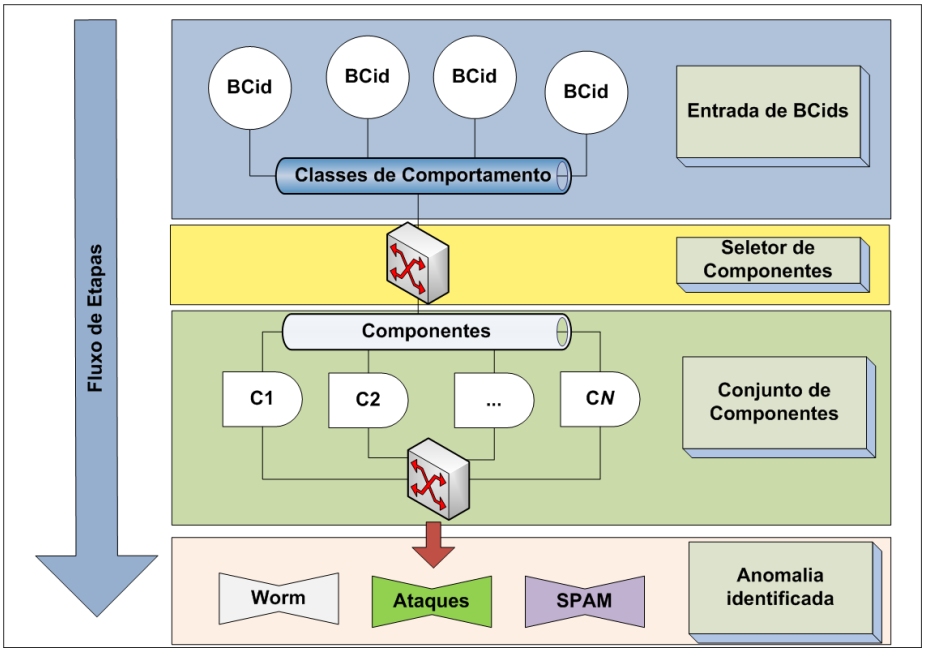


Figura 4.6: Processo de correlação e interpretação das classes de comportamento.

Para ilustrar como algumas anomalias de rede podem ser observadas a partir de características do padrão de comportamento, a Tabela 4.3 apresenta uma representação para as classes BC_3 e BC_{21}. A Tabela 4.3 é composta pela BC (Classe de Comportamento), o vetor de incerteza relativa (RU – *Relative Uncertainty*) usado no cálculo da BC, uma representação

gráfica do padrão de comportamento usando *graphlet* [20], indicativo das possíveis anomalias, e um breve comentário das características.

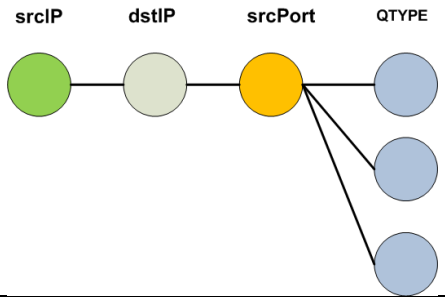
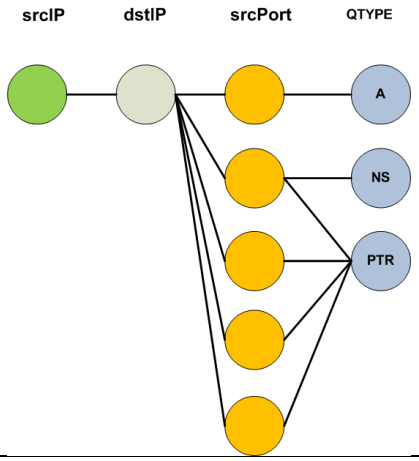
BCs	**Vetor RU**	**Representação em *graphlet***	**Possíveis Anomalias**	**Características**
BC=3	[0,1,0]		PTR-Scan, Força Bruta SSH,	Endereço IP de origem utilizando o registro PTR para varrer endereços de rede. Variação dos registros A, NS e PTR, sendo o PTR o mais freqüente no tráfego.
BC=21	[2,1,0]		*Mass-mailing worms*, spam	Nós de rede com o registro MX superior em relação ao tipo A, apresentando padrões de comportamento semelhantes.

Tabela 4.3: Características de anomalias de rede com base no padrão de comportamento.

Considere um ataque de *PTR-Scan*, observando o grupo chave endereço IP de origem em relação às dimensões correlacionadas. O comportamento desse tipo de ataque pode ser representado da seguinte forma: *srcIP* fixo (e.g. 200.132.45.6), porta de origem fixa (e.g., srcPort = 3200), endereço IP de destino fixo (e.g., dstIP = 200.160.0.10) e média dispersão de registro (QTYPE = **). Este trabalho usa três níveis discretos para indicar o grau de dispersão dos elementos com relação à dimensão analisada, representado da seguinte forma: * (baixa dispersão), ** (média dispersão) e *** (alta dispersão) conforme Seção 2.2. Logo, essa anomalia de rede poderá ser encontrada na classe de comportamento BC_3, isto é, o vetor de RU equivale [0,1,0]. Entretanto, outros tipos de anomalias também podem ser encontrados nessa classe de comportamento como, por exemplo, ataques de força bruta contra os serviços de SSH.

Para tornar mais claro o entendimento, considere outro exemplo de anomalia onde o endereço IP de origem é observado em relação aos demais recursos de tráfego. Considere ainda que os valores assumidos para porta de origem (srcPort = ***), sendo *** alta dispersão de valores, endereço IP de destino (dstIP = 200.160.0.10) e tipo de registro de recurso (QTYPE = **), sendo ** média dispersão de registros de recurso. Portanto, com base nos valores assumidos em cada dimensão correlacionada, essa anomalia pode ser encontrada na classe de comportamento BC_{21}, ou seja, o vetor de incerteza relativa assume [2,1,0]. Tal comportamento pode caracterizar ataques de propagação de mensagem em massa [31].

4.7.1 Inspeção do comportamento identificado

A segunda fase da interpretação da classe de comportamento é baseada na inspeção das características de cada classe. Essas características são consideradas pelo seletor de componentes para acionar os componentes que auxiliarão o processo detecção de anomalias. O seletor de componentes é baseados em regras estáticas do tipo "*if-the-else*", então para cada classe de comportamento, um conjunto de componentes é executado. Para tornar mais claro o entendimento, considere o grupo chave endereço IP de origem na classe de comportamento BC_3, cujo vetor de RU equivale [0,1,0]. O primeiro componente iniciado pelo seletor é o procedimento que valida à natureza do IP de origem observado. Essa validação permite identificar se o endereço é um servidor de nomes, servidor de e-mail, cliente de internet banda larga ou um nó de rede com serviços de internet. Detectado um cliente de internet banda larga, por exemplo, outro componente é executado para validar se o endereço está cadastrado em uma lista negra. Além disso, as consultas emitidas por esse hospedeiro também são confrontadas em listas negras, caso o tipo da consulta seja do tipo PTR. O componente que calcula a distribuição de consultas por tipo de registros enviado por esse cliente de internet banda larga é acionado. Esse último componente permite entender o padrão de consultas solicitadas por esse nó, como, por exemplo, o registro mais freqüente requisitado. Com bases nessas informações do nó observado é possível inferir o tipo de anomalia praticada.

Por exemplo, a Figura 4.7 ilustra a estrutura de regras estáticas que são utilizadas para detectar anomalias. Se o seletor de componentes recebe como entrada uma classe de comportamento BC_3 então um componente é executado, o grupo endereço IP de origem é

observado e se esse endereço for um cliente de internet banda larga, outros componentes são iniciados para indicar a anomalia de rede.

```
if BC3
then
   executa componente_1;
   if srcIP == Cliente ADSL
   then
      executa compomente_2 and componente_3;
   end if
end if
```

Figura 4.7: Regras estáticas utilizadas para classificar anomalias de rede pertencentes à classe de comportamento BC_3.

Componente	Descrição	Nome	Linguagem	Objetivo
C1	Recebe um endereço IP como parâmetro. Procura pelo endereço reverso.	*GetIPInformation*	*Perl* e *Shell Script*	Identificar a natureza do nó de rede.
C2	Recebe um endereço IP ou o nome da consulta do PTR e valida em um conjunto de listas negras.	*CheckRBL*	*Perl*	Validar se o endereço está cadastrado em uma lista negra.
C3	Recebe um endereço IP e contabiliza na base de dados o total de registro de recurso enviado pelo endereço.	*CalcDistByQType*	C	Calcular a distribuição de freqüência por tipo de registro de recurso por endereço informado.
C4	Abre a carga útil do pacote com base no endereço IP informado.	*DeepInspection*	C e *Perl*	Correlacionar informações como o tempo de vida do pacote, entropia da porta de origem e ID da consulta DNS.
C5	Calcula a Entropia de Shannon dos tipos de registros de recursos com base no endereço IP informado	*ShannonByQType*	C	Identificar o padrão de comportamento em ataques de rede. Permite observar os bits necessários para representar uma informação repetida.

Tabela 4.4: Lista de componentes utilizados para inspecionar as classes de comportamento por anomalia de rede.

Para ilustrar alguns dos componentes utilizados nessa fase de inspeção das classes de comportamento, a Tabela 4.4 apresenta a descrição, linguagem e o objetivo de cada

componente. É importante ressaltar que nessa fase, alguns componentes são correlacionados a partir das informações obtidas do componente anterior. Tal relação permite confirmar o comportamento anômalo detectado.

5 Análise de Desempenho

"A raiz de todos os males é o egoísmo."

Madre Teresa de Calcutá

Este Capítulo apresenta os resultados obtidos conforme o emprego da metodologia proposta descrita no Capítulo 4. Primeiro, uma análise passiva do tráfego é realizada visando caracterizar a composição do tráfego de dados analisado e identificar possíveis comportamentos anômalos. Em seguida, a visão geral das classes de comportamentos é demonstrada e como os componentes indicados no Capítulo 4 são utilizados para detectar anomalias na rede. Finalmente, cada classe de comportamento mais relevante para esse trabalho é apresentada. Nessas classes é possível observar diferentes tipos de anomalias de rede que utilizam o tráfego DNS e prejudicam o funcionamento da Internet.

5.1 Base de dados

A metodologia proposta é avaliada usando uma base de dados com registros de tráfego real cedida pelo OARC (*Operations, Analysis, and Research Center*) [9], que mantém em sua infraestrutura, os registros de tráfego coletados durante o projeto DITL 2008. O projeto DITL é uma ação colaborativa entre grandes servidores de nomes distribuídos ao redor do mundo. Cada servidor que participa do evento coleta, durante os dias determinados, o tráfego DNS de modo passivo. No Brasil, participaram do projeto cinco (*{a-e}.dns.br*) dos seis servidores de nome com autoridade pelo domínio *.br*.

O tráfego de rede capturado durante os dias do evento representa em média 5.4 bilhões de consultas, correspondendo a um volume equivalente a 230 GB de tráfego de rede compactado. A Tabela 5.1 apresenta um resumo da base de dados utilizada neste trabalho.

Por razões de segurança e privacidade, esses dados são armazenados e mantidos em servidores do Centro de Pesquisa OARC, sendo que todo o processamento e análise de dados

devem ser realizados dentro da própria infraestrutura de servidores disponibilizados pelo OARC.

DITL 2008		
Dia do Monitoramento	18 de Março de 2008	19 de Março de 2008
Instâncias envolvidas	{a-e}.dns.br	{a-e}.dns.br
Total de horas	24h	24h
Hora de início	00:00 UTC (+0000)	00:00 UTC (+0000)
Hora de término	23:59:59.999 UTC (+0000)	23:59:59.999 UTC (+0000)
Quantidade de pacotes	2.7 Bilhões	2.6 Bilhões

Tabela 5.1: Informações sobre base de dados do projeto DITL 2008.

A arquitetura oferecida pelo OARC dispõe de um sistema de arquivo que mantém o tráfego DNS coletado pelos servidores de cada país em diretórios. Cada diretório é composto por um conjunto de arquivos referente ao tráfego coletado no intervalo de uma hora. Por exemplo, o arquivo *20080319050000.pcap.gz*, localizado no diretório da instância *a.dns.br*, representa a coleta do tráfego referente ao dia 19 de março de 2008 que tem início às 05h e se estende até às 05h e 59min. No entanto, é importante destacar, nesse exemplo, que esse tráfego corresponde somente as consultas destinadas a instância *a.dns.br*, portando, para investigar o tráfego processado pela instância *c.dns.br*, é necessário localizar os arquivos dentro do diretório *c.dns.br*.

5.2 Metodologia de análise dos dados

A análise dos dados segue a metodologia proposta no Capítulo 4. Esta metodologia está codificada em uma ferramenta denominada DICA-DNS (Detecção, Identificação e Classificação de Anomalias através da análise do tráfego DNS). Esta ferramenta opera da seguinte forma: Primeiro, cada instância com autoridade pelo domínio .br tem o seu tráfego processado pela ferramenta DICA-DNS nos servidores da OARC. O ambiente de análise fornecido é um servidor FreeBSD, versão 6.4, com dois processadores AMD *Opteron*(tm) *Processo*r 846 e 32GB de memória principal. É importante ressaltar que esse ambiente é compartilhado com outros pesquisadores interessados na base de dados DITL.

Segundo, a leitura dos arquivos envolve o processo de agregação de fluxos de dados. Esses fluxos temporariamente ficam locados na memória principal até que o tempo (leitura

online) ou tamanho do arquivo (leitura *offline*) tenha terminado. Cada fluxo novo ou finalizado é armazenado em um arquivo texto. Um fluxo de dados é considerado válido durante o tempo máximo de 300 segundos. Esta quantidade de tempo tem sido utilizada com eficiência em [7] e [8].

Terceiro, finalizado o processo de agregação de fluxos de dados, o arquivo texto informado como parâmetro inicial é processado para que os grupos mais significativos sejam extraídos.

Quarto, o processo de classificação de classe de comportamento utiliza os grupos mais significativos encontrados pelo algoritmo na fase anterior. As classes de comportamento são encontradas a partir do cálculo da incerteza relativa.

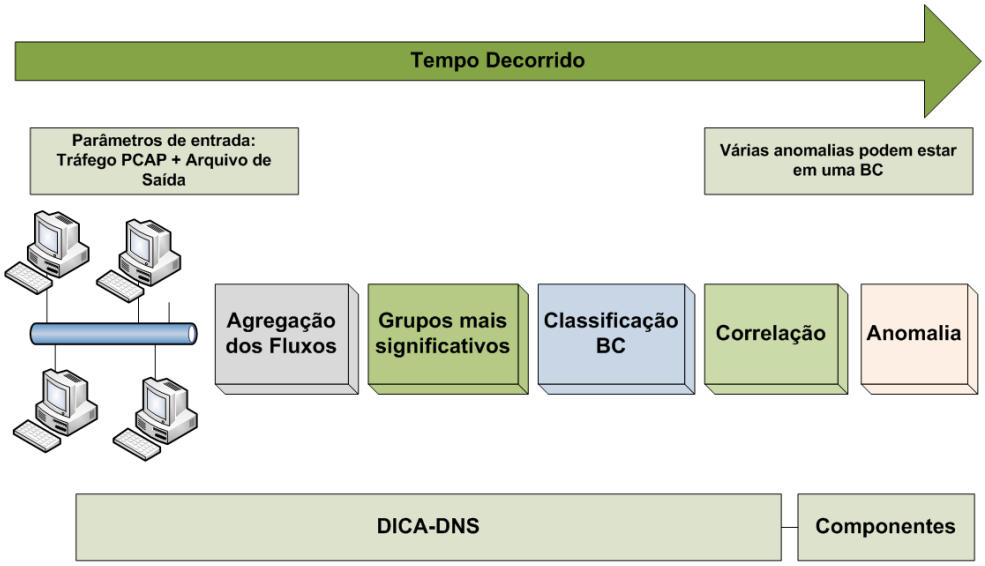


Figura 5.1: Resumo do processo de detecção, identificação e classificação de anomalias através do tráfego DNS.

Por último, finalizada a classificação das classes de comportamento, a ferramenta DICA-DNS aciona alguns componentes de inspeção com base nos resultados das classes. O objetivo dos componentes de inspeção é realizar inspeções no tráfego de dados que auxiliem no processo de interpretação dos comportamentos observados em cada classe. A Figura 5.1 sumariza o processo de detecção, identificação e classificação de anomalias através do tráfego DNS.

A ferramenta DICA-DNS foi codificada em linguagem C para as principais fases da metodologia e, os componentes de inspeção usados na correlação de informações, foram desenvolvidos nas linguagens *Perl* e *ShellScript*. A opção por componentes independentes é um ponto positivo da metodologia proposta, pois novos componentes ou novas regras podem ser incorporados na metodologia sem impactar nos demais processos (agregação de fluxos, extração dos grupos mais significativos e classificação dos grupos).

5.3 Resultados da análise dos dados

Esta seção apresenta os resultados da análise do tráfego de dados em duas etapas. Primeiro, uma análise passiva da base de dados utilizada é realizada visando caracterizar o comportamento do tráfego. Tal análise é importante, pois possibilita a compreensão da dinâmica de comunicação entre os nós na rede. Segundo, uma análise dos dados é realizada empregando a metodologia proposta neste trabalho. Os resultados obtidos demonstram que a metodologia proposta é capaz de detectar anomalias de rede pela análise do padrão de comportamento.

5.3.1 Caracterização passiva do tráfego DNS

Esta seção reporta a medição do tráfego DNS de forma passiva para caracterizar a composição do tráfego de dados analisado e identificar alguns comportamentos anômalos no tráfego. Em uma análise inicial foi possível observar resultados diferentes de outros trabalhos já realizados sobre medição do tráfego DNS [44].

O volume de tráfego estudado é uma fonte rica de informações referente aos comportamentos produzidos intencionalmente ou não. Por exemplo, consultas intencionais são produzidas por atividades maliciosas na rede como varredura de nós ou comunicação entre máquinas infestadas por aplicações de código malicioso. Entretanto, consultas não intencional decorrem de problemas de configuração da zona de domínio DNS ou consultas aos domínios de primeiro nível não reconhecidos pela IANA [71].

5.3.1.1 Distribuição de consultas por tipo de registro de recurso

Para ilustrar como as consultas DNS estão distribuídas na base de dados, a Figura 5.2 apresenta a distribuição de consultas por tipo de registro de recurso recebidas pelos servidores

de nome com autoridade pelo domínio .br. As consultas do tipo PTR, utilizadas para informar o nome de um domínio a partir de um endereço IP, são as principais solicitações utilizando esse recurso no tráfego, correspondendo a 42,91% do total de consultas. Por outro lado, a fração de consultas do tipo A, que corresponde aos registros que mapeiam nomes de máquinas para endereços IP dos hospedeiros, representa 30,29% do total de tráfego analisado, sendo que sua fração permanece relativamente estável durante dois dias do evento. O terceiro tipo de registro mais observado é o MX com 15,65% do total de registros. Este tipo de registro de recurso indica uma lista de servidores que devem receber e-mails para esse domínio. Os registros de recursos restantes são consultas do tipo AAAA e A6 que indicam consultas DNS buscando domínios utilizando endereços IP na versão 6. Estes recursos representam em média 6% e 2% do total, respectivamente. O tipo SVR, empregado para consultar serviços ou protocolo da rede, apesar da baixa taxa de participação no tráfego, correspondendo em média a 0,03% do total, também contribuem para a poluição com o DNS com consultas inválidas.

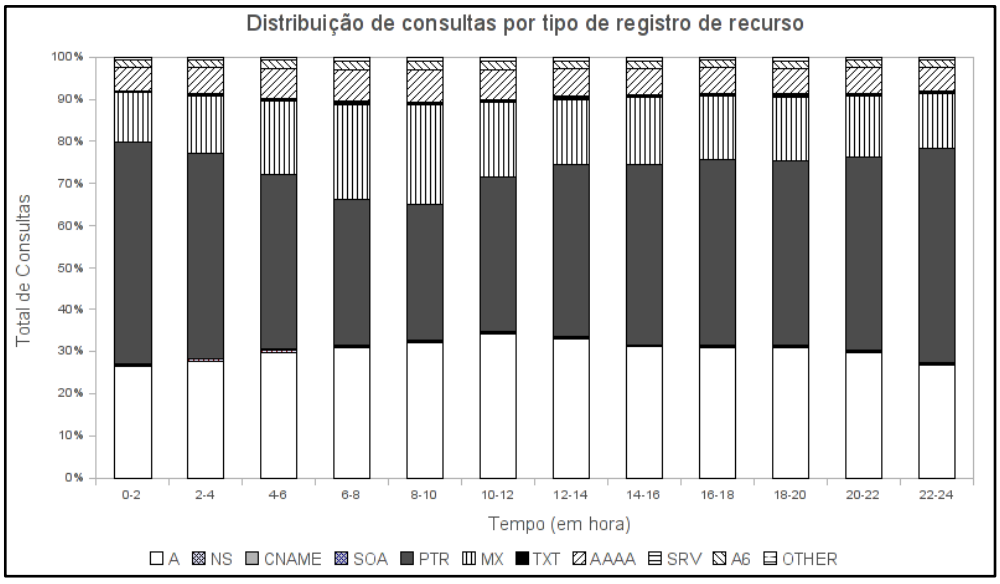


Figura 5.2: Distribuição de consultas por tipo de registro de recurso dos servidores {a-e}.dns.br.

Os resultados apresentados na distribuição de consultas relevam comportamento diferente dos resultados obtidos em [39] [43] e [5]. Nesses trabalhos, a distribuição de consultas por tipo de registro demonstra que, o recurso do tipo A, é o recurso mais casual no tráfego DNS devido ao acesso de páginas na Internet, assim como, algumas soluções anti-spam e resoluções DNS padrão. Entretanto, o resultado ilustrado na Figura 5.2, demonstra o

registro do tipo PTR sendo o mais freqüente no tráfego, portanto, este trabalho realiza uma investigação mais detalhada para encontrar as possíveis fontes que produzem tal resultado.

5.3.1.2 Determinando a validade das consultas

Detectar poluição no tráfego DNS é uma atividade desafiadora porque algumas aplicações, apesar de não obterem respostas válidas, ainda continuam funcionando. Em [39] é apresentado o processo de classificação de consultas inválidas através da utilização de uma lista que indica a categoria da qual uma pergunta pode ser definida. Essa lista também é aplicada em outros trabalhos de classificação e monitoramento do tráfego, como em [41] e [5].

Os resultados nas subseções seguintes aplicam algumas definições dessa lista, como nomes de domínio compostos de TLD inválido, nomes de consultas com o caractere "_" e consultas para os endereços definidos na RFC 1918. No entanto, outras métricas são utilizadas para detectar anomalias de rede, como identificação das principais fontes de registros do tipo PTR e classificação das consultas mais freqüentes no tráfego.

5.3.1.3 Consultas com o domínio de primeiro nível inválido

Interromper o tráfego DNS com o domínio de primeiro nível (TLD) inválido é uma tarefa de grande complexidade, devido à arquitetura atual do sistema DNS não oferecer nenhum mecanismo que diferencie entre domínios válidos e inválidos. Portanto, detectar, identificar e classificar fontes que degradam e consomem recursos da Internet, a partir de consultas não desejadas, é uma atividade importante para que os serviços de tradução de nome continuem em funcionamento.

Este trabalho considera como domínios de primeiro nível inválidos aqueles cujo sufixo não possuem o domínio .br. Além disso, domínios não reconhecidos pela IANA também são caracterizados nesta classe de anomalias.

Para ilustrar como a quantidade de consultas com o domínio de primeiro nível inválido pode degradar o serviço de resolução de nomes na Internet, a Tabela 5.2 apresenta os cincos domínios de primeiro nível mais consultados. É importante considerar que servidores de nome com autoridade pelo domínio *.br* apenas possuem informações ou referências dos servidores de domínio que estão em baixo da sua hierarquia de nomes. Nesse contexto, é possível

identificar atividades anômalas enviando consultas repetidas para os domínios apresentados na tabela ilustrada. Por exemplo, consultas para o domínio de primeiro nível *.org* podem ser registradas devido a problemas de configuração em soluções anti-SPAM. Além disso, solicitações para domínios *.gif, .jpg, .css* e *.js* denotam consultas mal formadas por erro de digitação ou implementações defeituosas em navegadores de Internet.

Domínio de Primeiro Nível	Quantidade Observada	Total em %
.com	1205130	30,72%
.net	828596	21,12%
.org	677228	17,26%
.gif / .js / .css / .jpg	910761	23,22%
.NULO	300000	7,64%
Total	3921715	100%

Tabela 5.2: Domínios de primeiro nível observados na base de dados DITL 2008.

5.3.1.4 Consultas inválidas que iniciam com "_"

Serviço de Rede	Nome da Consulta DNS
LDAP	_ldap._tcp.<Identificador>.<Domínio DNS>
Kerberos	_kerberos._tcp.dc._<Domínio DNS>

Tabela 5.3: Exemplo de consultas inválidas que iniciam com o caractere "_" encontradas nos servidores {a-e}.dns.br.

Problemas de configuração em aplicações podem produzir consultas que alcançam os servidores Raiz e consomem recursos disponíveis da infraestrutura de Internet desnecessariamente. Solicitações DNS iniciando com o caractere "_" são produzidas por clientes ou servidores de rede que fazem parte de um domínio do *Microsoft Windows Active Directory* [72]. Estas consultas utilizam o registro de recurso do tipo SRV para encontrar um serviço de rede disponível. A Tabela 5.3 apresenta o conteúdo dessas consultas anômalas e o serviço de rede utilizado.

Tais consultas representam 9.1 milhões das solicitações analisadas, o que corresponde a 0.35% do total do tráfego. Requisições desta natureza desperdiçam recursos da infraestrutura da Internet e, portanto, poderiam ser mitigadas a partir de simples práticas como adoção de

políticas que estabeleçam atualizações do sistema operacional como, por exemplo, através da aplicação de *patches* de atualização do produto.

5.3.1.5 Consultas inválidas com o caractere '@'

Consultas DNS com o caractere '@' não estão relacionadas como solicitações inválidas Entretanto, este trabalho sugere que essas consultas sejam consideradas como inválidas, uma vez que o símbolo '@', na configuração de zona de DNS, pode ser utilizado para identificar o endereço de correio eletrônico do administrador daquela zona [73]. Além disso, considerando o contexto de navegação em páginas de Internet, o caractere '@' é empregado para distinguir entre o nome de um usuário e o domínio de rede. Logo, é possível inferir que consultas apresentando essas características são originadas por problemas de configuração nas zonas de domínio, soluções de combate a mensagens não solicitadas (SPAM) ou erros de digitação de endereços de Internet.

Exemplo	Nome da Consulta
1	<Usuário>@<Domínio>
2	<Endereço IP>.<Usuário>@<Domínio>
3	<Endereço IP>@<Domínio>
4	@<Domínio>

Tabela 5.4: Consultas DNS que utilizam o caractere '@'.

Para tornar o entendimento mais claro, a Tabela 5.4 apresenta exemplos de requisições utilizando o caractere '@' encontradas nos servidores responsáveis pelo domínio .br . O exemplo 1 um demonstra consultas que buscam determinado usuário de domínio. No exemplo seguinte, as consultas são formadas por um endereço IP, nome do usuário e domínio. O endereço IP observado não representa o endereço IP que originou a requisição DNS ou o domínio consultado. Na verdade, os resultados obtidos denotam que esses endereços IPs, na grande maioria, são de clientes de Internet banda larga, portanto, é razoável assumir que servidores de correio eletrônico e servidores de nomes válidos estejam consultando se o <Endereço IP> pertence ao domínio. No terceiro exemplo somente endereços IP invertidos nome de domínio são observados e, finalmente, o quarto exemplo apresenta o caractere @ e um domínio.

Com base nos resultados da análise passiva é possível inferir que essas consultas sejam validações para mitigar o recebimento de mensagens não solicitadas. Por exemplo, considere os registros ilustrados na Tabela 5.5. Nessa tabela estão representados os endereços IP de origem e o nome da consulta DNS. Se o endereço IP *41.c.b.200* for extraído do nome da consulta e avaliado em listas negras, será possível constatar que esse endereço está cadastrado nesse tipo de base de dados, portanto, é possível inferir que o IP *41.c.b.200* esteja praticando atividades maliciosas. Essa observação também ocorre para os demais exemplos ilustrados e registrados durante análise do tráfego. É importante ressaltar que o endereço IP localizado no nome da consulta, não reflete o domínio do IP de origem.

Endereço IP de Origem	**Nome da Consulta DNS**
201.b.c.196	41.c.b.200.elencotreinamento@<Domínio>.com.br.
	6.c.b.200.envio@<Domínio>.com.br.
	37.c.b.201.@<Domínio>.com.br.
208.b.c.8	100.c.b.201.negociacaofinasa@<Domínio>.com.br
	mst.@<Domíno>.com.br
	192. c.b.200.cartao@<Domínio>.com.br

Tabela 5.5: Exemplo de consultas DNS utilizando o caractere '@'.

O total de solicitações com essa característica representa em média 4.2 milhões de consultas, isto é, 0.16% do total do tráfego observado.

5.4 Caracterização do tráfego DNS pela análise das classes de comportamento

Para ilustrar a visão geral das classes de comportamento, a Tabela 5.6 apresenta o total de nós classificados em cada classe de comportamento do grupo chave endereço IP de origem (*srcIP*). Por exemplo, o cálculo da classe BC_3 denota o comportamento de alta concentração para porta de origem, média dispersão para o tipo de registro de recurso e alta concentração para o endereço IP de destino, isto é, $BC_{id} = [0.3^2 + 1.3^1 + 0.3^0] = 3$, para maiores detalhes consulte a seção 4.7. Isso significa que padrões de comportamento que apresentam essa característica no tráfego, serão encontrados na classe BC_3.

Para tornar o entendimento mais claro, considere um provedor de Internet banda larga, uma anomalia de rede pode utilizar o registro de recurso do tipo NS e A para encontrar um

servidor de nomes com autoridade pelo serviço de Internet. Em seguida, o registro reverso (PTR) é manipulado para reconhecer possíveis clientes ativos nessa rede (*PTR-Scan*). Portanto, é possível assumir que algumas classes de comportamento registram com mais freqüência determinados registros de recursos considerando o padrão de comunicação entre os nós na rede.

Classe de comportamento para o grupo chave *srcIP*	**Servidores de nomes**				
	a.dns.br	**b.dns.br**	**c.dns.br**	**d.dns.br**	**e.dns.br**
BC=0	2243	2552	2589	2161	2765
BC=3	26028	30612	31484	29051	31695
BC=6	1921	2055	2221	1888	1676
BC=18	2462	2842	2731	2223	2814
BC=21	64380	76236	78190	69022	76683
BC=24	6822	7195	7796	6792	7973

Tabela 5.6: geral das classes de comportamento.

É importante ressaltar que as classes de comportamento apenas indicam os grupos mais significativos em termo de distribuição de domínio, portanto, para inferir atividades anômalas no tráfego de rede, é necessário analisar o padrão de comunicação considerando os recursos de tráfego correlacionado em cada classe. Tal análise utiliza os componentes descritos na Seção 4.7.1. As subseções a seguir apresentam as classes de comportamento e as métricas utilizadas para inferir padrões de comunicação anômalos.

5.4.1 Classe de comportamento: BC_0

A composição da classe de comportamento BC_0 envolve baixa entropia para as dimensões RU(X), RU(Y) e RU(Z), isto é, alta concentração para os recursos de tráfego correlacionados. Para tornar o entendimento mais claro, considere a observação do grupo chave endereço IP de origem (*srcIP*) em relação às dimensões porta de origem (*srcPort*), tipo da consulta (*QTYPE*) e endereço IP de destino (*dstIP*). O resultado nesta classe pode ser interpretado pelo comportamento de alta concentração para (*srcPort*), baixa entropia para o (*QTYPE*) e alta concentração para o endereço IP de destino.

Para a base de dados investigada, algumas anomalias foram detectadas nessa classe de comportamento. Por exemplo, ataques de reconhecimento de rede utilizando o tipo de registro

de recurso PTR, máquinas que compõem redes controladas por atacantes (botnets) consultando outros computadores infectados por aplicações maliciosas, ataques de reconhecimento de servidores de correio eletrônico e ataques de força bruta distribuído contra o serviço do SSH. A classe BC_0 é composta, na sua grande maioria, por servidores de correio eletrônico, *proxy*, servidores nome de domínios válidos e robôs indexadores de páginas de Internet. Tais servidores utilizam o tipo de registro do tipo A para localizar, por exemplo, endereços de sites ou servidores de nome. Para encontrar a natureza do nó presente nessa classe, o componente *GetIPInformation* é utilizado.

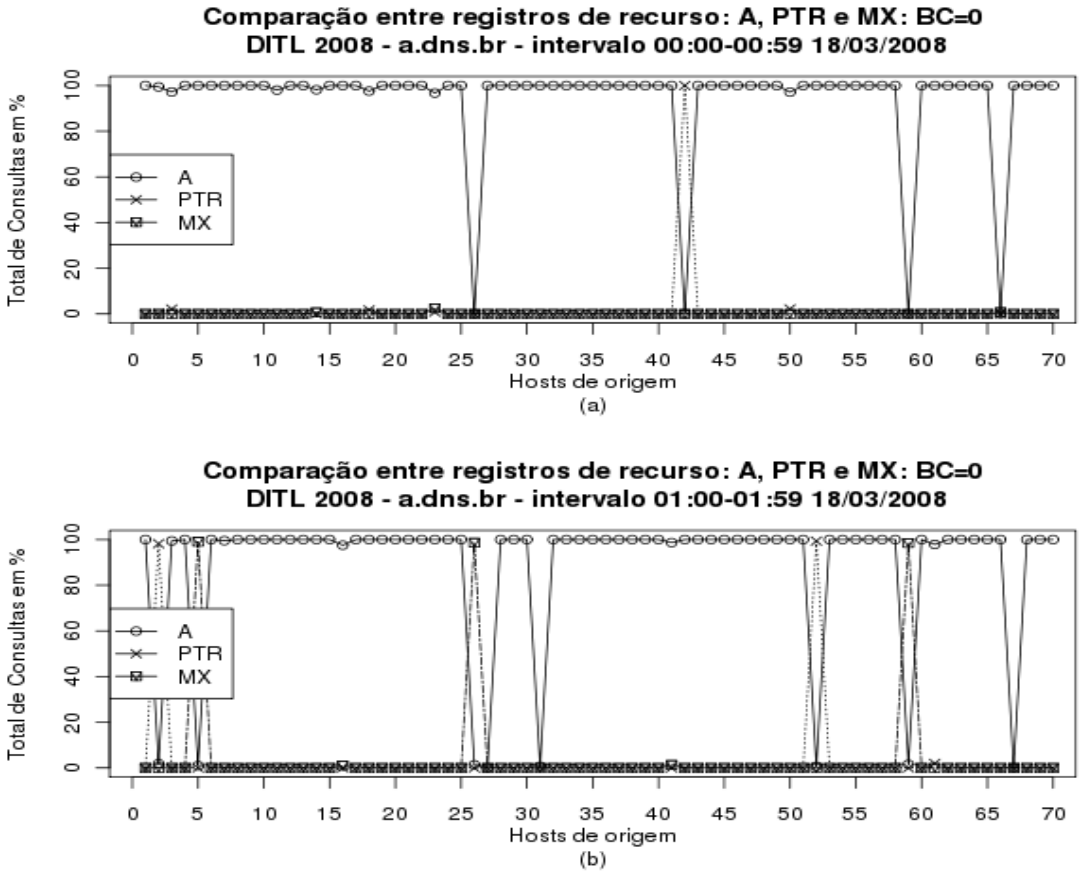


Figura 5.3: Comparação dos tipos de registro de recurso A, PTR e MX para classe BC=0 do grupo chave (srcIP) nos servidores a.dns.br (a) e b.dns.br (b).

A Figura 5.3 ilustra a comparação dos tipos de registro de recurso A, PTR e MX do servidor *a.dns.br* durante o intervalo 00:00-00:59 e 01:00-01:59 da manhã do dia 18 de março de 2008. Tal ilustração é resultado do componente que calcula a distribuição de freqüência por tipo de registro de recurso, *CalcDistByQType*. Em média, a classe BC_0 registra, em cada

intervalo de uma hora, 70 endereços IP de origem. Para os demais servidores, o comportamento é semelhante ao ilustrado na Figura 5.3. Em outras palavras, o registro do tipo A é contabilizado com mais freqüência nesta classe, como ilustra a Figura 5.3(a) e Figura 5.3(b).

Em [5] o registro do tipo A é observado com mais freqüência no tráfego devido ao acesso de endereços de Internet e, portanto, a composição deste tráfego possui maior incidência para esse tipo de registro. De acordo com [43], grandes porções de ocorrência do registro do tipo A são resultados de soluções de combate às mensagens não solicitadas em massa (SPAM). Além disso, técnicas evasivas de reconhecimento em listas negras, também podem influenciar a freqüência desse registro no tráfego. Máquinas que participam de redes controladas por criminosos realizam consultas em listas negras para validar se o seu endereço ou de outra máquina controlada, não está cadastrado na base de dados.

Além da grande concentração do registro de recurso do tipo A, outros registros também são encontrados nesta classe. Para elucidar esta situação, considere os eventos nos gráficos onde o registro de recurso do tipo A não é observado. Esses eventos representam anomalias de rede apresentando alta concentração para outros registros de recurso. Por exemplo, o nó 43 ilustrado na Figura 5.3 (a) é um cliente de rede de banda larga de um provedor de Internet da Rússia. Esse nó utiliza o registro reverso (PTR) para reconhecer endereços ativos (ataque do tipo *PTR-Scan*).

Para obter informações sobre um endereço IP como, por exemplo, país de origem e nome de domínio totalmente qualificado (FQDN), o componente de inspeção *GetIPInformation* deve ser utilizado. Esse componente é útil no processo de detecção de anomalias, pois, seu resultado permite ao seletor de componentes acionar outros componentes de inspeção para identificar anomalias de rede. Por exemplo, se um cliente de rede banda larga é registrado no tráfego, o componente de inspeção *CheckRBL*, o qual valida se o endereço IP ou os endereços consultados por ele, estão cadastrados em listas negras, é utilizado para validar se o nó observado apresenta comportamento malicioso.

Para o exemplo do nó 43 ilustrado na Figura 5.3(a), tanto o endereço IP de origem quanto os endereços consultados, estão presentes em algum tipo de lista negra. Um dado interessante dessa análise é que 73% dos endereços consultados, isto é, localizados no nome da consulta DNS, estão em listas negras. Esse comportamento permite inferir que o atacante

estava recrutando novos computadores com vulnerabilidades ou ainda, esta estratégia é uma validação do estado das máquinas que participam da rede controlada pelo atacante.

Outro exemplo de anomalia encontrada na classe BC_0 são ataques de dicionário distribuído ou força bruta contra servidores que oferecem o serviço SSH. Para tornar o entendimento mais claro, considere o nó 02 e 53 ilustrados na Figura 5.3(b). Diferente de um ataque de reconhecimento de rede, onde os endereços no nome da consulta são únicos, essa atividade apresenta alta freqüência para um conjunto de endereços consultados. Em outras palavras, o endereço IP de origem, que possui o serviço de SSH, é atacado por diferentes endereços constantemente. A Tabela 5.7 apresenta o número de consultas realizadas pelo endereço IP de origem e a freqüência em que o nome da consulta é registrada. Cada IP de origem consulta em média 16 vezes o nome da consulta por minuto, sendo que esses endereços estão 88% cadastrados em listas negras.

IP de Origem	**Consultas Realizadas**	**Freqüência observada**	
		Quantidade	**Nome da Consulta**
Host 2	325	107	100.b.c.d.in-addr.arpa.
		23	92 b.c.d.in-addr.arpa.
		15	159. b.c.d.in-addr.arpa.
		10	70.b.c.d.in-addr.arpa.
Host 53	472	20	28. b.c.d.in-addr.arpa.
		15	211. b.c.d.in-addr.arpa.
		13	47. b.c.d.in-addr.arpa.
		12	60. b.c.d.in-addr.arpa.

Tabela 5.7: Ilustração de um ataque distribuído de força bruta contra servidores que oferecem o serviço de SSH.

5.4.2 Classe de comportamento: BC_3

A composição da classe de comportamento BC_3 envolve baixa entropia para as dimensões RU(X), RU(Z) e média dispersão para RU(Y). Para tornar o entendimento mais claro, considere a observação do grupo chave endereço IP de origem (*srcIP*) em relação às dimensões porta de origem (*srcPort*), tipo da consulta (*QTYPE*) e endereço IP de destino (*dstIP*). O resultado nesta classe pode ser interpretado pelo comportamento de alta

concentração para (*srcPort*), média dispersão para o (*QTYPE*) e alta concentração para o endereço IP de destino (*dstIP*).

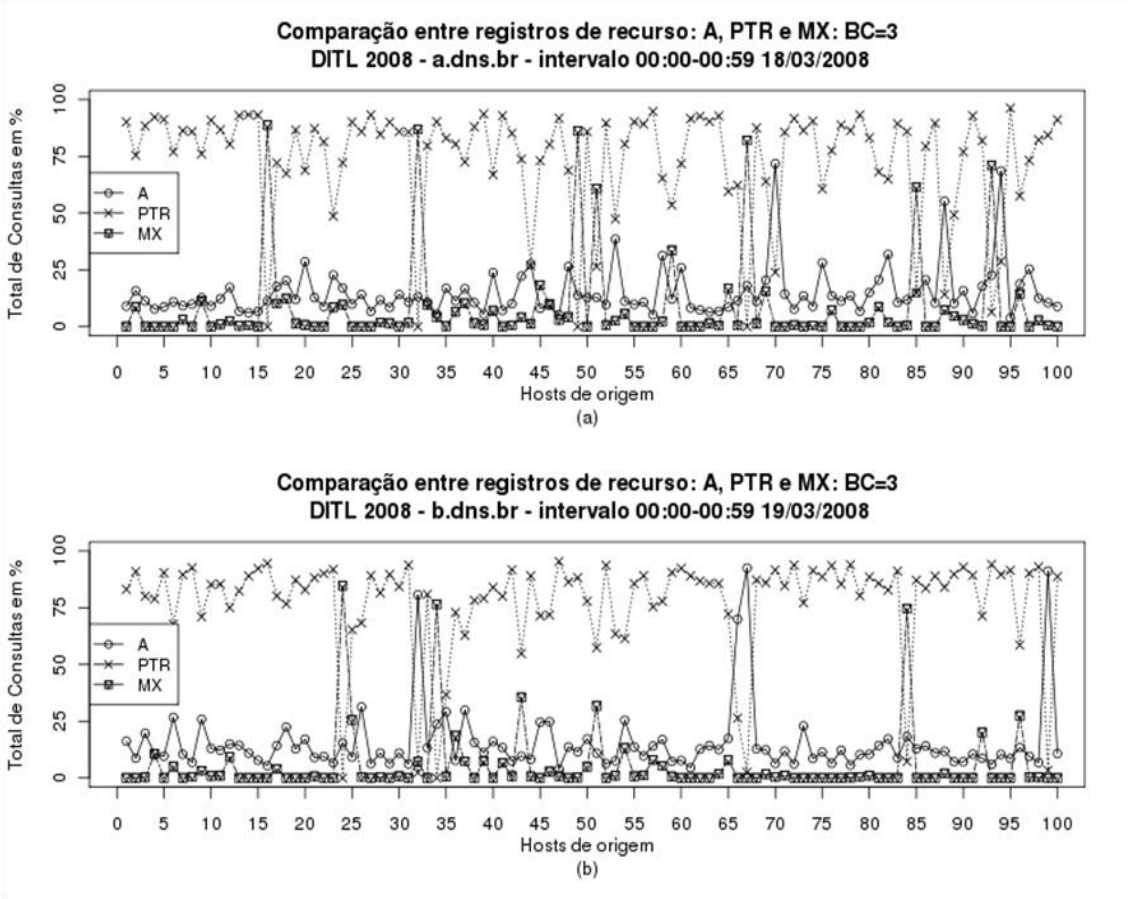


Figura 5.4: Comparação dos tipos de registro de recurso A, PTR e MX para classe BC=3 do grupo chave (srcIP) nos servidores a.dns.br (a) e b.dns.br (b).

Para a base de dados investigada, algumas anomalias podem ser observadas como ataques de reconhecimento de rede utilizando o tipo de registro de recurso PTR, envio de mensagens em massa enviadas por máquinas controladas por atacantes (*mass-mailing worms*), ataques de reconhecimento de servidores de correio eletrônico e ataques de força bruta distribuído contra o serviço do SSH.

A Figura 5.4 ilustra uma comparação dos tipos de registro de recurso A, PTR e MX dos servidores *a.dns.br* e *b.dns.br* durante o intervalo 00:00-00:59 da manhã, nos dias 18 e 19 de março de 2008, respectivamente. Esses registros foram escolhidos por serem os mais utilizados por anomalias de rede. Por razões visuais, o gráfico da Figura 5.4 apresenta apenas

a análise gerada pelo componente *CalcDistByQType* de 100 endereços IP. Como descrito na seção 4.7.1, este componente calcula a distribuição de consultas por tipo de registro de recurso.

Em média, a classe BC_3 registra, no intervalo de uma hora, 800 endereços IP de origem. Como mostrado na Figura 5.4, o registro do tipo PTR é encontrado com maior freqüência, entre 80-90% do total do tráfego analisado.

O registro reverso é utilizado tanto por aplicações válidas quanto atividades maliciosas. Para aplicações válidas, o registro do tipo PTR, por exemplo, pode ser utilizado para mitigar o recebimento de mensagens não autorizadas. Enquanto que, em atividades maliciosas, o registro pode ser empregado em ataques de reconhecimento de rede e ataques distribuídos de força bruta contra servidores que oferecem o serviço de SSH. Desta forma, a classe de comportamento BC_3 é uma fonte importante de comportamentos que empregam, com média freqüência, determinado registro de recurso.

Para ilustrar características de tráfego anômalo nesta classe, considere a Figura 5.5 onde são demonstrados os 30 primeiros endereços IP de origem observado na instância *a.dns.br*. Esses endereços são uma visualização ampliada dos nós da Figura 5.4(a). A ilustração na Figura 5.5(a) é resultado da análise do componente *ShannonByQType*, o qual calcula a entropia simples de *Shannon* para cada tipo de registro de recurso com base no endereço IP. Nesse gráfico é possível observar ataques de *mass-mailing worms* quando a entropia dos registros A e PTR está diminuindo. O entendimento desse comportamento foi obtido através dos componentes *GetIPInformation* e *CheckRBL*. O primeiro identifica a natureza o endereço IP de origem e o segundo, consulta os endereços IPs informados em listas negras. Por exemplo, os nós {1,3-5,8, 10, 13-15,21,25 e 27} são servidores de correio eletrônico consultando o endereço reverso de clientes de Internet banda larga. De acordo com a análise do componente *CheckRBL*, 87% dos clientes consultados por esses servidores estão cadastrados em listas negras.

Realizando uma comparação entre os gráficos das Figuras 5.5(a) e 5.5(b) é possível observar que, quando o registro possui alta freqüência relativa no tráfego menor é a quantidade de bits necessários para representar essa informação. Por outro lado, quando o

evento é único, ou possui baixa freqüência, são necessários mais bits para representá-lo, isso acontece devido à falta de certeza para aquele evento.

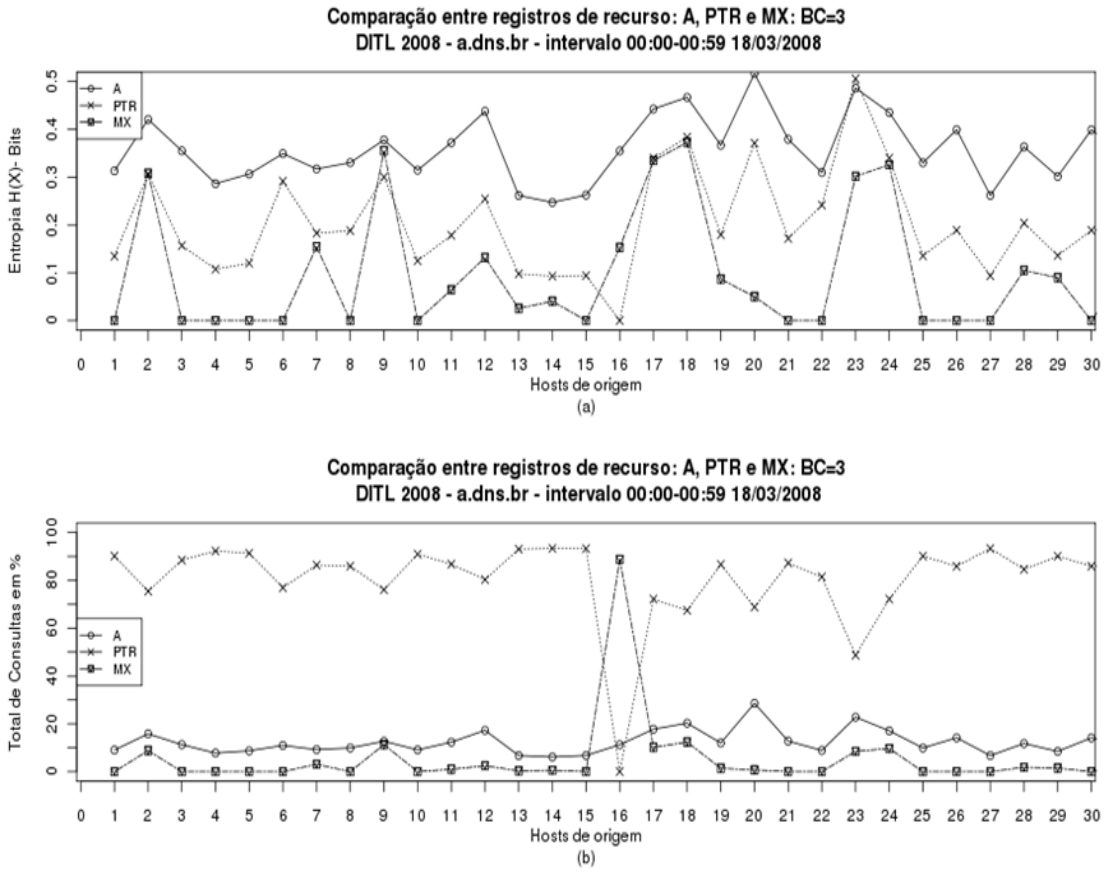


Figura 5.5: Ilustração dos 30 primeiros nós de rede, considerando a entropia (a) por tipo de registro de recurso e a distribuição de freqüência (b) por tipo de registro de recurso.

Outro exemplo de anomalia de rede que pode ser observada na classe de comportamento BC_3, são ataques de hospedeiros infectados por programas de envio de mensagem em massa. É importante ressaltar uma diferença entre esse tipo de anomalia e o exemplo de comportamento supracitado. Pois, enquanto que os servidores de correio eletrônico estão sendo atacado por *spam bots*, o exemplo ilustrado a seguir, denota o comportamento e características de hospedeiros que utilizam o tráfego DNS para enviar mensagens não solicitadas em massa.

Por exemplo, para evadir sistemas detectores de intruso (IDS), programas de código malicioso de propagação em massa utilizam aplicações SMTP independentes e, portanto,

antes de enviar uma mensagem eletrônica, um nó de rede comprometido por esse tipo de anomalia realiza consultas do tipo MX para encontrar possíveis alvos de ataque.

Para tornar o entendimento mais claro, considere o nó de rede 16 ilustrado na Figura 5.5(b). Utilizando o componente *DeepInspection* que extrai informações da carga útil do pacote, é possível observar algumas características peculiares. Por exemplo, este nó de rede emite solicitações DNS com o bit ativo de recursividade desejada (RD) e suas requisições, possuem o identificador das consultas (ID) com valores inferiores a 255. É importante ressaltar que o campo do identificador de consultas oferece 2^{16} possibilidades distintas.

As características do nó 16 ilustrado na Figura 5.5(b), além de outros hospedeiros dentro desta mesma classe, apresentam comportamento similar quanto ao envio de consultas do tipo MX e bit RD ativo. Por essa razão, é possível inferir que esses nós de rede fazem parte de redes controladas por criminosos sem autorização. Os resultados obtidos nesse trabalho também apresentam características similares ao trabalho de [27].

Para ilustrar o cenário exposto, uma abordagem mais detalhada sobre o comportamento de nós de rede infectados por *mass-mailing worms* é avaliada a seguir. Além do bit RD ativo e valores inferiores a 255 para o identificador da pergunta DNS (ID), outras características semelhantes são compartilhadas por hospedeiros nesta classe. Por exemplo, mesmo valor para o TTL do cabeçalho IP, endereços de clientes de Internet banda larga cadastrados em listas negras, bits do campo de fragmentação do cabeçalho IP não ativos e o total de consultas do tipo de registro de recurso MX acima de 75% das consultas enviadas.

É importante observar que, para a base de dados investigada, quando um nó de rede apresenta as características supracitadas, é possível assumir que são máquinas comprometidas por alguma aplicação de código malicioso. Este trabalho assume que essas características de padrão de comunicação na rede, sejam uma assinatura de máquinas infectadas por programas para envio de mensagem em massa.

Neste contexto, para validar essa hipótese, a Figura 5.6 demonstra uma comparação entre 200 para os tipos de registro A e MX da instância *a.dns.br* Figura 5.6(a) e c.dns.br Figura 5.6(b) durante o intervalo 00:00-00:59 do dia 18 de março de 2008, respectivamente. Para as demais instâncias com autoridade do domínio .br, o comportamento para essa base de

dados também é observado. A linha horizontal vermelha é o limiar que separa os nós com essa característica de comportamento.

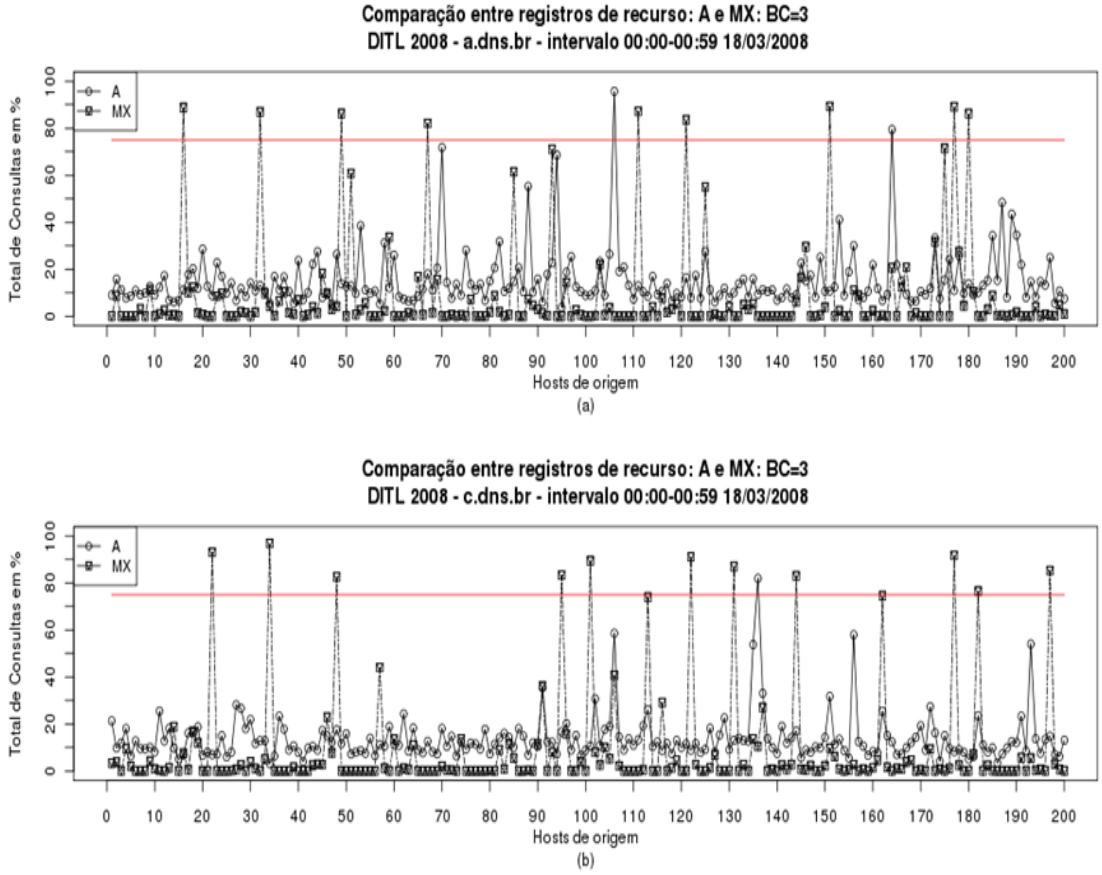


Figura 5.6: Comparação entre registro de recurso do tipo A e MX para as instâncias a.dns.br (a) e c.dns.br (b) durante o intervalo 00:00-00:59 do dia 18 de março de 2008.

Os hospedeiros localizados acima do limitar observado confirmam a hipótese proposta, em outras palavras, cada nó de rede teve sua carga válida aferida através dos componentes *GetIPInformation*, *CheckRBL*, *CalcDistByQType* e *DeepInspection*. O total de hospedeiros infectados que foram atendidos pela instância *a.dns.br*, nesse exemplo, representam 4,5% do total de nós ilustrados na Figura 5.6(a). Enquanto que, 5,5% indicam os hospedeiros comprometidos na Figura 5.6 (b).

A hipótese assumida equivale a 100% de precisão de nós de rede infestados por aplicações para envio de mensagens em massa. Isso significa que cada nó é avaliado individualmente quanto à carga útil da pergunta, assim como o padrão de comportamento na

rede. Anomalias como mass-mailing worms são registradas durante toda análise da base de dados do projeto DITL 2008. A Figura 5.7 ilustra o total de nós de rede, do grupo chave (srcIP), atendidos pela instância a.dns.br, no horário 01h Figura 5.7(a) e 02h Figura 5.7(b) da manhã, do dia 18 de março de 2008, respectivamente.

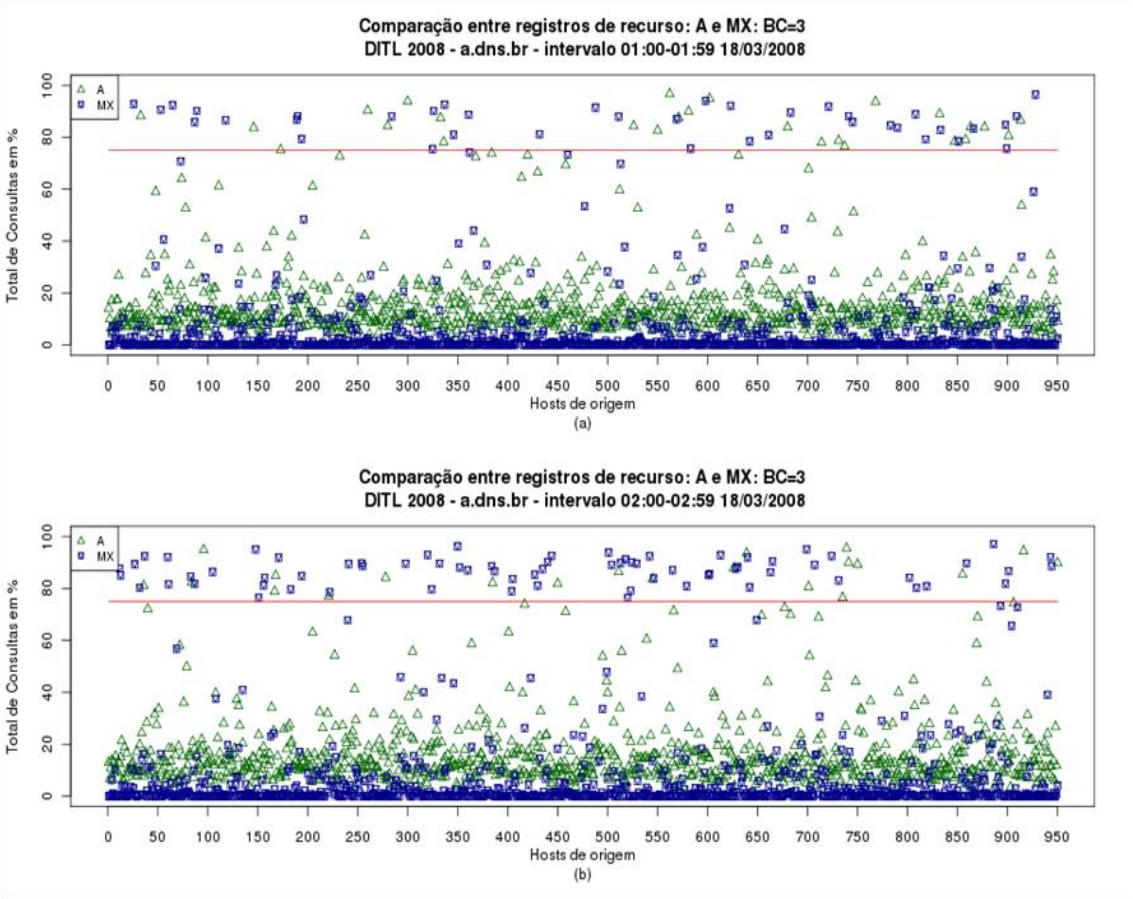


Figura 5.7: Comparação entre os registros de recurso A e MX para a instância a.dns.br durante o intervalo 01:00-01:59(a) e 02:00-02:59(b).

No exemplo ilustrado a distribuição de consultas dos registros A e MX são aferidas. A grande concentração de nós observados abaixo de 40% denota comportamento considerado como normal no tráfego. Por outro lado, pontos quadrados acima do limite estabelecido representam nós de rede utilizados como *mass-mailing worms*. Os pontos triangulares, por sua vez, indicam a presença de servidores de nomes recursivos consultando, além do registro do tipo A, outros recursos como A6, NS, AAAA e TXT.Os valores ilustrados nas Figura 5.7(a) e Figura 5.7(b) estão sumarizados na Tabela 5.8. Neste resumo, pode ser observada a quantidade total de consultas registradas em cada intervalo de hora, o total de consultas

apresentada em cada gráfico, a quantidade de anomalias monitoradas considerando o limite de 75% e o total em porcentagem de anomalias detectas.

Intervalo Observado	Qtd. Consultas	Qtd. Ilustrada	Qtd. Anomalias a 75%	Total de Anomalias em %
01:00:00-01:59:59	985	950	35	3.68
02:00:00-02:59:59	957	950	64	6.73
Total	1942	1900	99	5.21

Tabela 5.8: Sumário dos comportamentos observados na instância a.dns.br durante o intervalo de 1h e 02h do dia 18 de março de 2008.

É importante ressaltar que o limite de 75% estipulado considera o resultado da análise passiva das características da classe de comportamento BC_3. Além disso, o valor estabelecido em 75% reflete 100% de certeza para essa anomalia. Ou seja, valores próximos ao limite de 75%, não indicam com precisão se o endereço IP de origem é uma atividade maliciosa. Entretanto, é possível que essa anomalia seja identificada em outras classes de comportamentos, apresentando outras características como, por exemplo, alta variação para a porta de origem.

5.4.3 Classe de comportamento: BC_6

A composição da classe de comportamento BC_6 envolve baixa entropia para as dimensões RU(X), RU(Z) e alta dispersão para RU(Y). Para tornar o entendimento mais claro, considere a observação do grupo chave endereço IP de origem (*srcIP*) em relação às dimensões porta de origem (*srcPort*), tipo da consulta (*QTYPE*) e endereço IP de destino (*dstIP*). O resultado nesta classe pode ser interpretado pelo comportamento de alta concentração para (*srcPort*), alta dispersão para o (*QTYPE*) e alta concentração para o endereço IP de destino (*dstIP*).

Um ponto peculiar dessa classe de comportamento é que os registros de recursos estão melhores dispersos em relação à distribuição de freqüência. Em outras palavras, enquanto nas classes de comportamento BC_0 e BC_3 existe a predominância de um registro de recurso, a classe BC_6 mantém seus valores concentrados, abaixo de 60% do total de consultas enviadas. Essa observação pode ser constatada através do resultado do componente *CalcDistbyQType*. Os valores obtidos podem ser organizados em uma tabela de freqüência. Por exemplo, para

cada intervalo de hora analisada, a quantidade de registros que foram observados dentro do intervalo entre 0-20% são contabilizados. É importante destacar que esse intervalo de classe não considera o tipo de registro de recurso mais freqüente, na verdade, esses valores indicam apenas a quantidade de recursos contabilizados em cada intervalo de classe definido. A Tablea 5.9 sumariza essas observações para os registros de recurso A, PTR, MX, AAAA, TXT e ANY.

Intervalo em %	**Observações**
00 - 20	182
20 - 40	459
40 - 60	305
60 - 80	130
80 - 100	1

Tabela 5.9: Distribuição de freqüência dos registros de recurso A, PTR, MX, AAAA, TXT e ANY.

Para a base de dados investigada, é possível identificar comportamentos maliciosos para o envio de mensagens não solicitadas em massa. Entretanto, deve ser considerado que existe um conjunto de variantes de vermes que podem apresentar características semelhantes no tráfego de rede, ou seja, utilizam o tráfego DNS para disseminar SPAM. Entretanto, a classificação (nome do verme) desse tipo de anomalia, não pode ser obtida apenas através da análise passiva do tráfego DNS, isso porque, muitos *worms*, para não serem identificados encontrados, encriptam seu código fonte [25]

A Figura 5.8 ilustra uma comparação dos tipos de registro de recurso A, PTR e MX dos servidores *a.dns.br* e *b.dns.br* durante o intervalo 00:00-00:59 da manhã, nos dias 18 e 19 de março de 2008, respectivamente. Os gráficos exibidos na Figura 8 são resultantes da análise do componente *CalcDistByQType*. Os registros A, PTR e MX são correlacionados devido a sua importância para atividades maliciosas no tráfego. Servidores de correio eletrônico podem ser observados sendo atacados por vermes de envio de mensagem em massa. Para tornar o entendimento mais claro, considere os nós {2,3} ilustrados na Figura 5.8(a) e {5,22} apresentados na Figura 5.8(b). Através do componente *GetIPInformation*, *CheckRBL*, *CalcDistByQType* e *DeepInspection*, é possível inferir que esses hospedeiros estão sob ataques de *mass-mailing worms*.

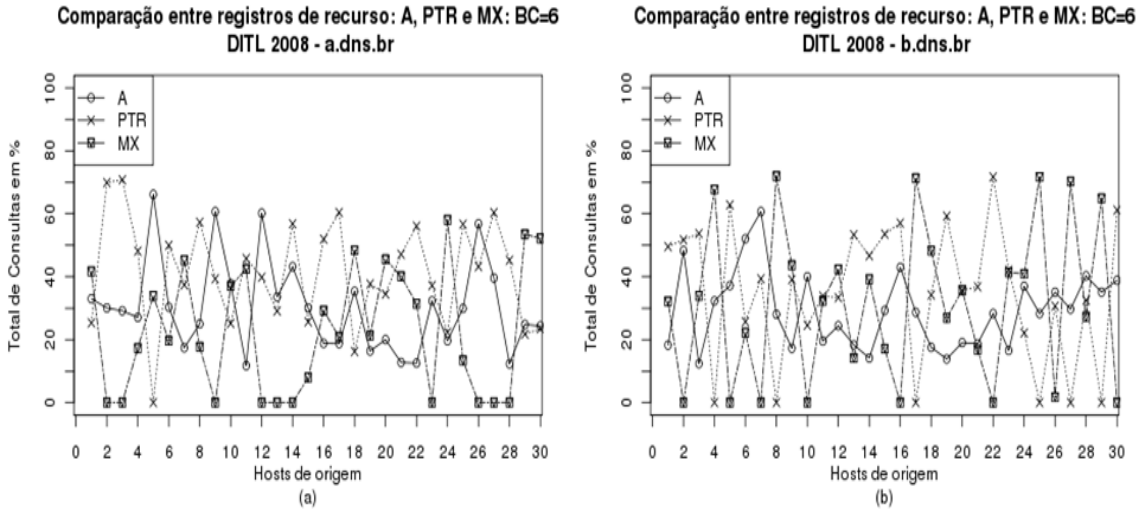


Figura 5.8: Comparação entre os registros de recurso A, PTR e MX para a instância a.dns.br(a) e b.dns.br(b) durante o intervalo 00:00-00:59.

Com base nas observações a partir da análise passiva da classe de comportamento BC_6, é possível estabelecer um limite que permite identificar atividades maliciosas no tráfego. Entretanto, é importante ressaltar que esse limite reflete as características da base de dados e, portanto, para outras bases, é possível que o limite seja diferente. Para tornar o entendimento mais claro, considere a ilustração da Figura 5.9, resultante do componente que calcula a distribuição de freqüência por tipo de registro de recurso. A partir das observações acima do limite estabelecido em 60%, é possível identificar algum tipo de anomalia de rede.

Por exemplo, considere as observações de consultas do tipo MX acima de 65%. Por meio dos componentes já mencionados, cada nó identificado foi avaliado individualmente e representam máquinas comprometidas por programas de código malicioso de mensagem em massa. Para validar essa hipótese, cada nó de rede teve sua carga válida investigada pelo componente que extrai informações da carga útil do pacote. O resultado obtido é semelhante ao comportamento da classe BC_3 para hospedeiros infectados por *mass-mailing worms*. Em outras palavras, estes endereços IP de origem possuem baixa entropia para o identificador da consulta DNS (ID), bit de recursividade desejada (RD) ativo e tempo de vida (TTL) do cabeçalho IP dentro do intervalo 45-46 ou 110-111. Além disso, esses endereços são clientes de Internet de banda larga, sendo 94% cadastrados em listas negras.

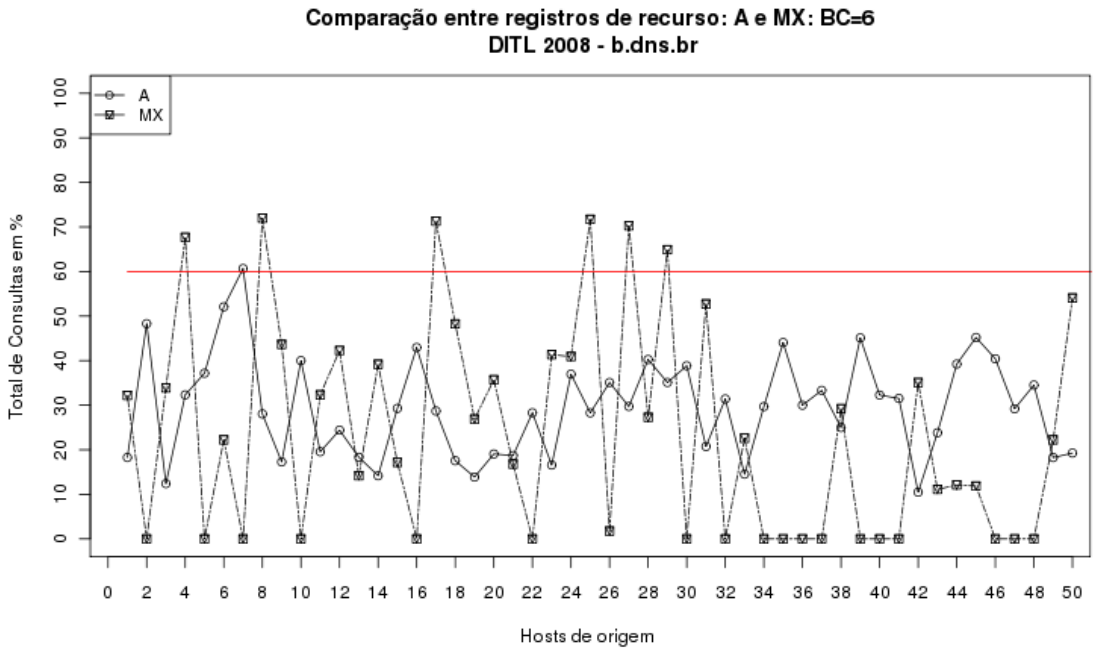


Figura 5.9: Comparação entre os registros de recurso A e MX para a instância b.dns.br durante o intervalo 00:00-00:59.

5.4.4 Classe de comportamento: BC_{21}

A composição da classe de comportamento BC_{21} envolve alta dispersão para dimensão RU(X), média dispersão para RU(Y) e alta concentração para RU(Z). Para tornar o entendimento mais claro, considere a observação do grupo chave endereço IP de origem (*srcIP*) em relação às dimensões porta de origem (*srcPort*), tipo da consulta (*QTYPE*) e endereço IP de destino (*dstIP*). O resultado nesta classe pode ser interpretado pelo comportamento de alta dispersão para (*srcPort*), média dispersão para o (*QTYPE*) e alta concentração para o endereço IP de destino (*dstIP*).

Para a base de dados investigada, é possível observar ataques de redes maliciosas para enviar mensagem não solicitada em massa. A classe de comportamento BC_{21} apresenta comportamento anômalo semelhante ao resultado observado em BC_3, considerando é claro, o padrão de comunicação dos nós de rede.

A Figura 5.10 ilustra uma comparação dos tipos de registro de recurso A, PTR e MX do servidor *a.dns.br* durante o intervalo 00:00-00:59 da manhã, do dia 18 de março de 2008. Os

registros A, PTR e MX são relevantes para demonstração, pois são mais utilizados por anomalias de rede. Por razões visuais o gráfico apresenta apenas a análise gerada pelo componente *CalcDistByQtype* para 200 endereços IP. Em média, a classe BC_{21} registra, em cada intervalo de hora, 1900 endereços IP de origem. Para os demais servidores, o comportamento é semelhante ao ilustrado na Figura 10. Nesta ilustração é possível observar dois grupos concentrados. O primeiro, referente ao registro PTR, está no intervalo entre 70-90% do tráfego. O segundo, registros A e MX, estão localizados abaixo dos 50%. Entretanto, é possível identificar consultas do tipo MX alcançando o primeiro grupo de concentração.

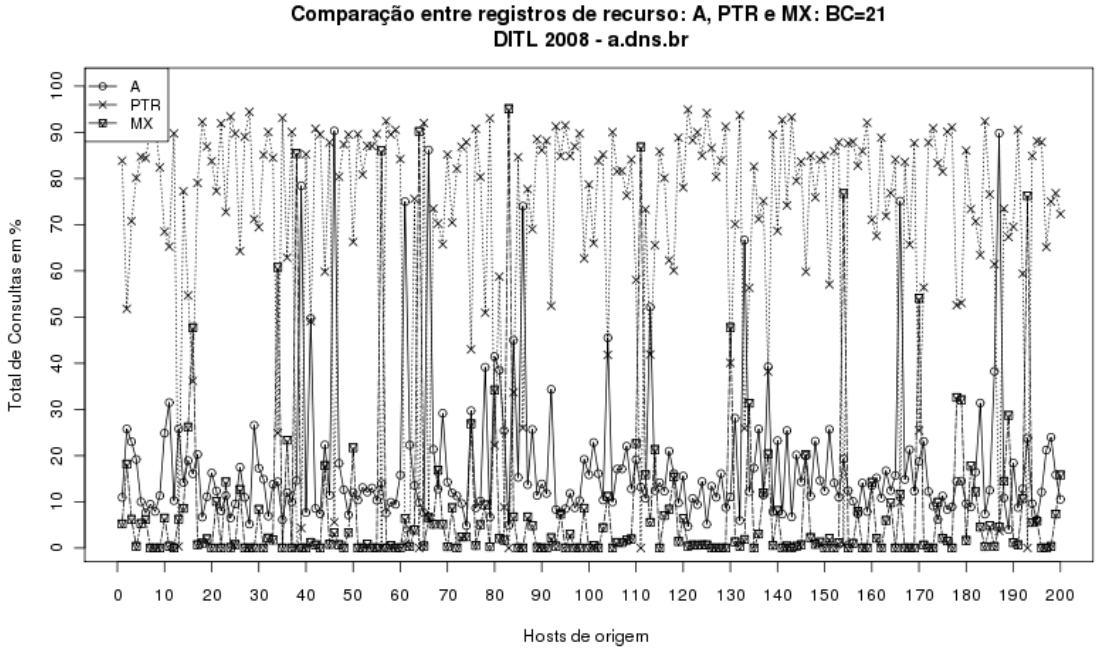


Figura 5.10: Comparação entre os registros de recurso A, PTR e MX para a instância *a.dns.br* durante o intervalo 00:00-00:59.

Avaliando o comportamento desses nós de rede através dos componentes *GetIPInformation*, *CheckRBL* e *DeepInspection* é possível assumir que esses hospedeiros compõem redes controladas por atacantes para enviar mensagens não solicitada em massa. Em outras palavras, se o limite de 75%, estipulado na classe BC_3, também for considerado nesta classe, essa validação corresponde a 100% de eficiência de detecção. No entanto, comparando o total de anomalias do tipo SPAM registradas nas classes BC_3 e BC_{21}, a classe BC_{21} apresenta quantidade inferior de anomalias desse tipo.

Por um lado, a classe BC_3 registra maior quantidade de máquinas enviando SPAM, por outro, a classe BC_{21} apresenta servidores de correio eletrônico recebendo ataques de redes controladas em maior freqüência. Essa conclusão pode ser aferida quando os tipos de registros de recurso A e PTR são comparados. Ambos os registros podem ser utilizados por filtros anti-spam. Por exemplo, o registro do tipo A é utilizado para validar se um endereço IP está cadastrado em lista negra. No entanto, o registro do tipo PTR pode ser utilizado por servidores de correio eletrônico, para confrontar se o endereço IP recebido na mensagem representa o domínio informado.

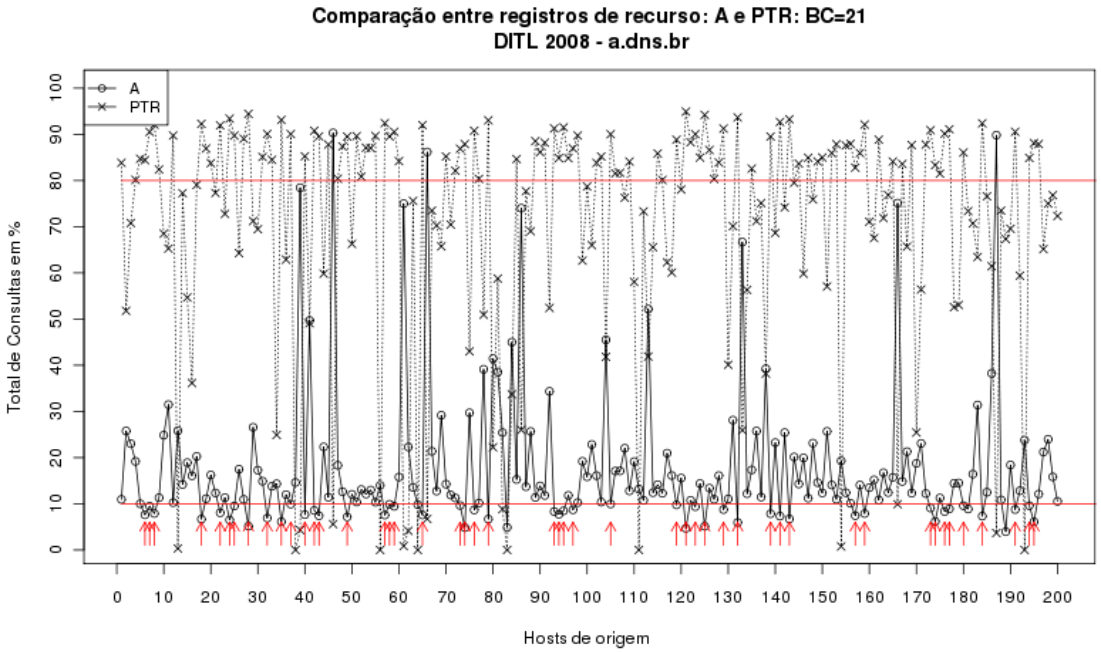


Figura 5.11: Comparação entre os registros A e PTR . Setas em vermelho indicam nós de rede sendo atacados por *mass-mailings worms*.

Para tornar o entendimento mais claro, considere a Figura 5.11, a qual apresenta a distribuição de consultas entre os registros de recurso A e PTR do grupo chave endereço IP de origem. Os resultados ilustrados no gráfico foram obtidos a partir do componente *CalcDistByQType*. A partir dos componentes *GetIPInformation DeepInspection* e *CheckRBL* é possível identificar servidores de correio eletrônico sendo atacados por máquinas controladas por atacante (*spam bots*). Essa constatação foi possível através da correlação entre esses componentes de inspeção e análise passiva do tráfego. Por exemplo, estabelecendo o limite para consultas do tipo PTR a 80% e consultas do tipo A em 10%, foi identificado um

padrão de ataques de *mass-mailing worms*. Em outras palavras, quando o registro PTR é observado em 80% das consultas e o tipo A com valores abaixo de 10% para servidores de correio eletrônico, essa anomalia pode ser confirmada.

Nesse contexto, é importante ressaltar que 96% dos endereços consultados por esses servidores são de clientes de Internet banda larga e, as consultas do tipo A, são realizadas para encontrar o servidor de nome com autoridade desses endereços IP. A Figura 5.11 ilustra 200 endereços IP de origem, sendo que 24% estão recebendo ataque de *mass-mailing worms*.

Um ponto positivo obtido durante o resultado dessa classe é que o tráfego da Internet tem sido utilizado para disseminação de mensagens não solicitas e prática de atividades maliciosas. Essa observação condiz com os resultados preliminares da base DITL realizada em [44].

5.4.5 Classe de comportamento: $\mathbf{BC_{24}}$

A classe de comportamento BC_{24} denota alta dispersão paras as dimensões RU(X) e RU(Y) e grande concentração para dimensão RU(Z). Para tornar o entendimento mais claro, considere a observação do grupo chave endereço IP de origem (*srcIP*) em relação às dimensões porta de origem (*srcPort*), tipo da consulta (*QTYPE*) e endereço IP de destino (*dstIP*). O resultado nesta classe pode ser interpretado pelo comportamento de alta dispersão para (*srcPort*), alta dispersão para o (*QTYPE*) e alta concentração para o endereço IP de destino (*dstIP*).

Para a base de dados investigada, é possível observar ataques de reconhecimento de rede através do registro reverso (*PTR-Scan*).

Com base nas observações a partir da análise passiva da classe de comportamento BC_{24}, é possível estabelecer um limite que permite identificar atividades maliciosas no tráfego. Entretanto, como já mencionado é importante ressaltar que esse limite reflete as características tanto da base de dados quanto da classe de comportamento. Para tornar o entendimento mais claro, a Figura 5.12 ilustra uma comparação dos tipos de registro de recurso A, PTR e MX do servidor *c.dns.br* durante o intervalo 00:00-00:59 da manhã, do dia 19 de março de 2008. Os registros A, PTR e MX são relevantes para demonstração, pois são mais utilizados por anomalias de rede. Por razões visuais o gráfico apresenta apenas a análise

gerada pelo *CalcDistByQType* de 100 endereços IP. Em média, a classe BC_{24} registra, em cada intervalo de hora, 120 endereços IP de origem. Para os demais servidores, o comportamento é semelhante ao ilustrado na Figura 5.12.

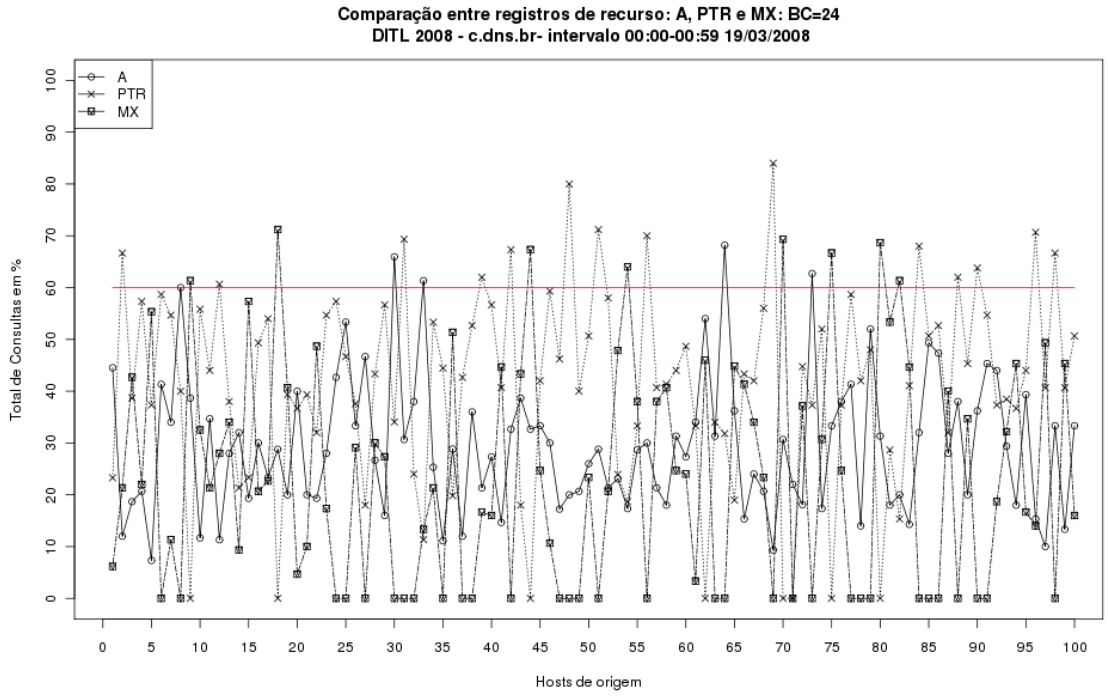


Figura 5.12: Comparação entre os registros de recurso A, PTR e MX para a instância *c.dns.br* durante o intervalo 00:00-00:59.

As observações acima do limite estabelecido em 60% apresentam comportamentos anômalos da rede. Por exemplo, o nó de origem 2 apresenta alta freqüência do registro PTR. Através do componente *DeepInspection* e *GetIPInformation* é possível aferir esse ataque de rede. É importante ressaltar que esse nó apresenta uma assinatura no processo de reconhecimento de rede através do registro reverso.

Para tornar o entendimento mais claro, considere a Tabela 5.10 que sumariza dois padrões de assinatura de reconhecimento de rede. O atacante estabelece a rede de ataque (A ou B) e o último octeto do endereço IP, sendo os demais octetos pseudo-aleatórios. Nessa assinatura de rede, o criminoso reconhece, por exemplo, os endereços da rede 90, em seguida, da rede 9, depois rede 80, e depois rede 8 e assim consecutivamente. Essa estratégia é interessante para evadir sistemas detectores de intrusos que aguardam por um padrão de varredura consecutivo. Nessa abordagem, os limites inferiores e superiores da rede, por

exemplo, 90 ou 9, são reconhecidos sem deixar uma assinatura visível. Esse padrão de varredura também é observado por [25]. Os nós de rede 47,48, 56 e 69 apresentam o mesmo padrão de reconhecimento de rede. Logo, é possível assumir que esses computadores façam parte da mesma rede controlada.

Padrão de Assinatura Rede A	**Padrão de Assinatura Rede B**
99.c.d.A.in-addr.arpa	88.c.d.B.in-addr.arpa
98.c.d.A.in-addr.arpa	88.c.d. B.in-addr.arpa
92.c.d.A.in-addr.arpa	88.c.d. B.in-addr.arpa
92.c.d.A.in-addr.arpa	87.c.d. B.in-addr.arpa
91.c.d.A.in-addr.arpa	87.c.d. B.in-addr.arpa
91.c.d.A.in-addr.arpa	87.c.d. B.in-addr.arpa
90.c.d.A.in-addr.arpa	87.c.d. B.in-addr.arpa
9.c.d.A.in-addr.arpa	84.c.d. B.in-addr.arpa
9.c.d.A.in-addr.arpa	82.c.d. B.in-addr.arpa

Tabela 5.10: Padrão de ataque de reconhecimento de rede

Considerações Finais

"A educação, se bem compreendida, é a chave do progresso moral."

Allan Kardec

5.5 Conclusões

O funcionamento correto do sistema de tradução de nomes é crucial para a operação de aplicações que dependem de suas funcionalidades como servidores de correio eletrônico, sistemas de mensagens eletrônicas e sites de comércio eletrônico. Devido à importância do DNS, este trabalho propôs e desenvolveu uma metodologia para detecção, identificação e classificação de anomalias através do tráfego DNS.

Os resultados mostraram que através de medidas da Teoria da Informação, mais especificamente, entropia de Shannon, combinada com recursos do tráfego de rede como endereço IP, portas de origem e informações da carga útil do sistema de tradução de nomes foi possível identificar atividades maliciosas no tráfego de rede. Essa identificação ocorreu devido ao padrão de comportamento das anomalias de rede que utilizam o tráfego DNS como ponto de partida para comprometer ou interromper os serviços de Internet.

A utilização da entropia no processo de detecção de anomalias é um ponto positivo, pois sua característica é a sumarização dos padrões concentrados ou dispersos de comunicação na rede. Em outras palavras, a entropia apenas indica alterações mais significantes e pouco comuns nos padrões de volume de tráfego, portanto, a quantidade de informação para análise é menor. Por isso, sua aplicação tem sido utilizada com freqüência em grandes volumes de tráfego.

A principal contribuição deste trabalho foi o emprego da entropia para detectar anomalias a partir do tráfego DNS. Enquanto que, outros trabalhos utilizam a entropia para agrupar fluxos de dados a partir da quíntupla bem conhecida, esta metodologia investigou quais recursos de tráfego (endereço IP, informação do DNS) poderiam ser melhores

explorados para identificar atividades maliciosas no tráfego. Desta forma, este trabalho contribuiu de forma significativa no processo de detecção de anomalias a partir da análise do tráfego DNS.

A partir de validações em registros de tráfego real, foi possível constatar a eficiência da metodologia proposta. Para as hipóteses assumidas durante o processo de pesquisa, os resultados demonstram 100% de detecção de anomalias sugeridas. Isso mostra que o processo investigatório, em relação quais recursos de tráfego seriam melhores aplicados nessa metodologia, foi satisfatório. Além disso, correlacionar componentes de inspeção observando os resultados das classes de comportamento permitiu alcançar esse nível de precisão.

Os resultados ainda confirmam as observações obtidas a partir do estudo preliminar da base de dados. O trabalho em [44] sugere que o alto índice do registro PTR é reflexos de atividades maliciosas no tráfego. Por exemplo, a metodologia proposta é capaz de detectar variações de comportamentos a partir de máquinas infectadas por aplicações de código malicioso. Um exemplo claro dessa observação são ataques de rede a partir de redes controladas contra servidores de correio eletrônico e máquinas infectadas realizando consultas do tipo MX para disseminar SPAM. Desta forma, os exemplos citados podem ser utilizados como base em outros estudos da base de dados DITL.

Outro ponto a ser destacado foi o modelo da arquitetura dos principais componentes deste trabalho, pois podem ser combinados por outros trabalhos de investigação de anomalias através do tráfego DNS. Isso significa que os componentes de correlação e inspeção das classes de comportamento são dinâmicos, desta forma, pesquisadores interessados em contribuir com novas métricas ou componentes de inspeção, podem agregar seu módulo sem afetar os principais componentes da metodologia.

Entretanto, alguns desafios encontrados durante o processo de pesquisa ainda precisam de investigações. Por exemplo, a base de dados utilizada é uma coleção de tráfego real obtida durante o projeto DITL de 2008, esse tráfego representa as consultas DNS que foram atendidas pelas instâncias com autoridade pelo domínio .br. Essa situação envolve algumas considerações. Primeiro, consultas DNS que são atendidas por servidores de nome com autoridade por código de países refletem apenas informações que estão abaixo da hierarquia daquele país. Segundo, servidores de nome de domínio de primeiro nível aplicam filtros de rede para mitigar algumas anomalias como consultas a partir de endereços IP forjado e

ataques de negação de serviço [41]. Terceiro, servidores de primeiro nível apenas respondem o caminho do servidor de nomes mais próximo, desta forma, ataques de amplificação de resposta DNS a partir desses servidores, não surtem efeito. E finalmente, apesar da base de dados possuir representação de tráfego de rede real, as consultas DNS refletem a realidade de cada instância individualmente, em outras palavras, o tráfego da instância *c.dns.br* possui apenas consultas para esse endereço. Nesse contexto, ataques de negação de serviço distribuído são mais difíceis de detectar.

Além disso, como as anomalias de rede também estão localizadas na carga útil da mensagem DNS, foi possível detectar durante os experimentos, que o nome da consulta DNS é relevante no processo de detecção de anomalias, no entanto, utilizar essa informação no processo de extração de grupos mais significativos é inviável devido ao volume de dados que é armazenado na tabela de espalhamento, e ainda, as características desse campo. Em outras palavras, o nome da consulta DNS não faz distinção entre caracteres maiúsculo e minúsculo, desta maneira, os exemplos *UFAM.com.br* e *ufam.Com.BR* alocariam espaço distintos na tabela de espalhamento, consumindo desnecessariamente recurso para análise.

Outra desvantagem da metodologia proposta é o tempo necessário utilizado pelos componentes de inspeção para correlacionar informações das classes de comportamento. Portanto, ainda que a metodologia proposta seja capaz de detectar anomalias de rede como ataques de redes controladas e propagação de mensagens não solicitada em massa, em análise em tempo real, a metodologia proposta não é viável.

5.6 Trabalhos Futuros

O valor assumido para o estimador ε, no cálculo da incerteza relativa, incorpora os conceitos utilizados pela proposta original do artigo [57]. No entanto, esses valores consideram a correlação e observação entre os nós de rede e serviço, enquanto que, este trabalho observa o tráfego DNS para identificar tráfego anômalo na rede. Por essa razão, é necessário um trabalho comparativo quanto à eficiência da metodologia proposta empregando outros valores para o estimador ε considerando o comportamento do tráfego DNS.

As classes de comportamento apresentadas refletem um conjunto de anomalias que são mais relevantes para o funcionamento do tráfego DNS. Por isso, é importante que novos ataques de rede sejam mapeados na base de dados DITL, resultando em novas classes de

comportamento. Além disso, é preciso avaliar a eficiência da metodologia proposta na base de dados DITL 2009 e 2010, desta forma, esses resultados servirão de análise para outros trabalhos comparativos.

Outro ponto a ser considerado como trabalho futuro foi à utilização de alguns recursos do tráfego DNS durante o processo de correlação. Por exemplo, o nome e o identificador da consulta foram correlacionados em alguns momentos para inferir tráfego anômalo na rede. Isso significa que novas dimensões podem ser adicionadas para indicar o grau de dispersão ou concentração no processo de extração dos grupos mais significativos.

Cada recurso de tráfego adicionado, entretanto, eleva o tempo computacional para extração dos grupos mais significativos. Desta forma, uma abordagem para contornar esse problema é a utilização de um conjunto de regras com base no recurso do tráfego observado. Por exemplo, se o tipo de registro de recurso do fluxo de dados correlacionado indicar o registro reverso (PTR), então as dimensões correlacionadas seriam endereço IP de origem e destino, nome da consulta (QNAME) e o tipo de registro. Essas características denotam algumas anomalias no tráfego como ataques de reconhecimento de rede, ataques de força bruta contra o serviço de SSH e clientes enviando mensagens não solicitadas em massa na rede.

Outra solução para identificar comportamento malicioso no tráfego de rede, é a partir da combinação de algoritmos de mineração de dados como redes bayesianas e técnicas de agrupamento como *k-means*. Esses algoritmos seriam empregados para substituir os módulos e métricas adicionais no processo de interpretação e correlação de dados. Em outras palavras, através do resultado em classes de comportamentos, esses algoritmos são empregados em um grupo de dados menor, pois a entropia indicou apenas os grupos mais significativos para análise. Métricas da Teoria da informação em combinação com algoritmos de mineração de dados é uma estratégia adotada por outros trabalhos [7]. Portanto, essa combinação é uma estratégia eficiente para detectar anomalias de rede.

5.7 Publicações

As publicações abaixo são resultado do processo de pesquisa durante a realização deste trabalho. A primeira delas foi aceita e as demais estão em processo de avaliação.

- Kaio Rafael de Souza Barbosa e Eduardo Souto, "Análise Passiva do Tráfego DNS da Interet Brasileira," *IX Simpósio Brasileiro em Segurança da Informação e de Sistemas Computacionais. SBSEG 2009*, pp. 203-216, Outubro 2009.
- Kaio Rafael de Souza Barbosa, Eduardo Souto, Ademir José de Carvalho Junior e Djamel Sadok, "An DNS-based Entropy Detection Anomaly Methodology". (Em avaliação)
- Kaio Rafael de Souza Barbosa e Eduardo Souto, "Análise Passiva do Domínio .BR: DITL 2009 e 2010". (Em avaliação)

Referências Bibliográficas

[1] Paul V. Mockapetris, Domain names - concepts and facilities, 1987.

[2] Jon Oberheide, Manish Karir e Z. Morley Mao, "Characterizing Dark DNS Behavior," DIMVA '07: *Proceedings of the 4th international conference on Detection of Intrusions and Malware, and Vulnerability Assessment*, pp. 140-156, 2007.

[3] Hyunsang Choi, Hanwoo Lee, Heejo Lee, e Hyogon Kim, "Botnet Detection by Monitoring Group Activities in DNS Traffic," *CIT '07: Proceedings of the 7th IEEE International Conference on Computer and Information Technology*, pp. 715-720, 2007.

[4] P. Kammas, T. Komninos, e Y. C. Stamatiou, "Modeling the Co-evololution DNS Worms and Anti-worms in IPv6 Networks," *IAS '09: Proceedings of the 2009 Fifth International Conference on Information Assurance and Security*, pp. 171-174, 2009.

[5] S. Castro, D. Wessels, M. Fomenkov, e K. Claffy, "A day at the root of the internet," *ACM SIGCOMM Computer Communication Review*, pp. 41-46, 2008.

[6] Claude E. Shannon e Warren Weaver, "A Mathematical Theory of Communication," *Bell System Technical Journal*, vol. 27, pp. 27:379-423, 623-656, Jul 1948.

[7] Anukool Lakhina, Mark Crovella, e Christophe Diot, "Mining anomalies using traffic feature distributions," *SIGCOMM Comput. Commun. Rev.*, vol. 35, pp. 217-228, 2005.

[8] Kuai Xu, Z. Zhang e S. Bhattacharyya, "Profiling internet backbone traffic: behavior models and applications," *Conference on Applications, Technologies, Architectures, and Protocols For Computer Communications* , 2005.

[9] DNS-OARC. (2009) DNS-OARC | The DNS Operations, Analysis, and Research Center. https://www.dns-oarc.net. Acesso em 14/04/2009.

[10] Douglas E. Comer, *Interligação em rede com TCP/IP*. Campus, 2007.

[11] Charles M. Kozierok, *The TCP/IP Guide*.: No Starch Press, 2005.

[12] Paul V. Mockapetris, Domain names - implementation and specification, 1987.

[13] Cricket Liu e Paul Albiz, *DNS and Bind*. Chicado: ORelly, 2007.

[14] James F. Kurose e K. W Ross, *Computer Networking: a Top-Down Approach. 5th.*: Addison-Wesley Publishing Company, 2009.

[15] Atkins, D. e R. Austein, Threat Analysis of the Domain Name System (DNS), 2004.

[16] Arends, R., R. Austein, M. Larson, D. Massey e S. Rose, DNS Security Introduction and Requirements, 2005.

[17] Suranjith Ariyapperuma e Chris J. Mitchell, "Security vulnerabilities in DNS and DNSSEC," *ARES '07: Proceedings of the The Second International Conference on Availability, Reliability and Security*, pp. 335-342, 2007.

[18] Steven M. Bellovin, "A Look Back at "Security Problems in the TCP/IP Protocol Suite"," *ACSAC '04: Proceedings of the 20th Annual Computer Security Applications Conference*, pp. 229-249, 2004.

[19] Boaz Tsaban, "Bernoulli numbers and the probability of a birthday surprise," *Discrete Appl. Math.*, vol. 127, pp. 657-663, 2003.

[20] Thomas Karagiannis, Konstantina Papagiannaki e Michalis Faloutsos, "BLINC: multilevel traffic classification in the dark," SIGCOMM '05*: Proceedings of the 2005 conference on Applications, technologies, architectures, and protocols for computer communications*, pp. 229-240, 2005.

[21] Paul V. Mockapetris, Sue Thomson, Y. Rekhter e Jim Bound, Dynamic Updates in the Domain Name System - DNS UPDATE, 1997.

[22] Fanglu Guo, Jiawu Chen e Tzi-cker Chiueh, "Spoof Detection for Preventing

DoS Attacks against DNS Servers," *Distributed Computing Systems, 2006. ICDCS 2006. 26th IEEE International Conference on*, pp. 37- 37, 2006.

[23] Paul V. Mockapetris, Gerry Sneeringer e Mark Schleifer. (2002) Events of 21-Oct-2002. Acesso em 10/07/2009.

[24] Changhua Sun, Bin Liu e Lei Shi, "Efficient and Low-Cost Hardware Defense Against DNS Amplification Attacks," *Global Telecommunications Conference, 2008. IEEE GLOBECOM 2008*, pp. 1-5, Novembro 2008.

[25] Jose Nazario, *Defense and Detection Strategies against Internet Worms*. Norwood, MA, USA: Artech House, Inc, 2003.

[26] Pin Ren, John Kristoff e Bruce Gooch, "Visualizing DNS traffic," VizSEC '06: *Proceedings of the 3rd international workshop on Visualization for computer security*, 2006.

[27] Keisuke Ishibashi et al., "Detecting mass-mailing worm infected hosts by mining DNS traffic dat," *In Proceedings of the 2005 ACM SIGCOMM Workshop on Mining Network Data*, pp. 159-164, Agosto 2005.

[28] Ruiqi Hu e Aloysius K. Mok, "Detecting Unknown Massive Mailing Viruses Using Proactive Methods," *In Proc. of the 7th Int'l Symposium on Recent Advances in Intrusion Detection*, 2004.

[29] David Whyte, P.C. Van Oorschot e Evangelos. Kranakis, "Addressing SMTP-Based Mass-Mailing Activity within Enterprise Networks," *Computer Security Applications Conference*, pp. 393-402, 2006.

[30] Ryuichi Matsuba e Kenichi Sugitani, "Indirect Detection of Mass Mailing Worm-Infected PC terminals for Learners," *Proceedings for the 3rd International Conference on Emerging Telecommunications Technologies and Applicatications (ICETA2004)*, pp. 233-237, 2004.

[31] Elias Levy, "The making of a spam zombie army. Dissecting the Sobig worms," *Security Privacy, IEEE*, pp. 58 - 59, 2003.

[32] Jian Zhang, Zhen-Hua Du e Wei Liu, "A Behavior-Based Detection Approach to Mass-Mailing Host," *Machine Learning and Cybernetics, 2007 International Conference on*, pp. 2140 - 2144 , Agosto 2007.

[33] Thomas M. Cover e Joy A. Thomas, Elements of Information Theory (Wiley Series in Telecommunications and Signal Processing). Wiley-Interscience, 2006.

[34] George J. Klir, Uncertainty and Information: Foundations of Generalized Information Theory.: Wiley-Interscience, 2005.

[35] Claude. E. Shannon e W Weaver, A Mathematical Theory of Communication., 1948.

[36] Aleksandr Yakovlevich Khinchin, Mathematical Foundations of Information Theory. 1957.

[37] Yakov Rekhter, B. Moskowitz, D. Karrenberg, G. J. de Groot e E. Lear, Address Allocation for Private Internets, 1996.

[38] Peter B. Danzig, Katia Obraczka e Anant Kumar, "An analysis of wide-area name server traffic: a study of the Internet Domain Name System," *SIGCOMM Comput. Commun. Rev.*, vol. 22, pp. 281-292, 1992.

[39] Duane Wessels e Marina Fomenkov, "That's a lot of packets," *in Proc. 2003 Passive and Active Measurements Workshop*, 2003.

[40] Duane Wessels, "Is your caching resolver polluting the internet?," *NetT '04: Proceedings of the ACM SIGCOMM workshop on Network troubleshooting*, pp. 271-276, 2004.

[41] N. Brownlee, K Claffy e E. Nemeth, "DNS Measurements at a Root Server," *Global Telecommunications Conference, 2001. GLOBECOM '01. IEEE*, pp. 1672-1676, 2001.

[42] CAIDA. (2008) A Day in the Life of the Internet (DITL).

[43] Bojan Zdrnja, Nevil Brownlee e Duane Wessels, "Passive Monitoring of DNS Anomalies," DIMVA '07: Proceedings of the 4th international conference on

Detection of Intrusions and Malware, 2007.

[44] Kaio Rafael de Souza Barbosa e Eduardo Souto, "Análise Passiva do Tráfego DNS da Interet Brasileira," *IX Simpósio Brasileiro em Segurança da Informação e de Sistemas Computacionais. SBSEG 2009*, pp. 203-216, Outubro 2009.

[45] CAIDA. (2002) Nameserver DoS Attack October 2002. "http://www.caida.org/projects/dns/dns-root-gtld/oct02dos.xml". Acesso em 21/08/2009

[46] Amit Klein. (2007) Trusteer. "http://www.trusteer.com/list-context/publications/bind-9-dns-cache-poisoning. Acesso em 12/07/2009

[47] Mark Santcroos e Olaf M. Kolkman. (2009) NLnet Labs. "http://www.nlnetlabs.nl/downloads/se-consult.pdf". Acesso em 12/07/2009

[48] David White, Evangelos Kranakis e P.C. van Oorschot, "DNS-based Detection of Scanning Worms in an Enterprise Network," *In Proceedings of the 12th Network and Distributed System Security Symposium*, Fevereiro 2005.

[49] Nikolaos Chatzis e Enric Pujol, "Email Worm Mitigation by Controlling the Name Server Response Rate," *Emerging Security Information, Systems, and Technologies, The International Conference on*, vol. II, pp. 139-145, 2008.

[50] Cynthia Wong, Stan Bielski, Jonathan M. McCune e Chenxi Wang, "A study of mass-mailing worms," *WORM '04: Proceedings of the 2004 ACM workshop on Rapid malcode*, pp. 1-10, 2004.

[51] Uwe Aickelin, Jamie Twycross, e Thomas Hesketh-Roberts, "Rule generalisation in intrusion detection systems using SNORT," *Int. J. Electron. Secur. Digit. Forensic*, pp. 101-116, 2007.

[52] Giovanni Vigna, William Robertson e Davide Balzarotti, "Testing network-based intrusion detection signatures using mutant exploits," *CCS '04: Proceedings of the 11th ACM conference on Computer and communications security*, pp. 21-30, 2004.

[53] Wenke Lee e Dong Xiang, "Information-Theoretic Measures for Anomaly Detection," *SP '01: Proceedings of the 2001 IEEE Symposium on Security and Privacy*, p. 130, 2001.

[54] Ricardo Villamarin-Salomon e J.C. Brustoloni, "Identifying Botnets Using Anomaly Detection Techniques Applied to DNS Traffic," *Consumer Communications and Networking Conference, 2008. CCNC 2008.*, pp. 476-481, Janeiro 2008.

[55] Ricardo Villamarin-Salomon e J.C. Brustoloni, "Bayesian bot detection based on DNS traffic similarity," *SAC '09: Proceedings of the 2009 ACM symposium on Applied Computing*, pp. 2035--2041, 2009.

[56] Andrew W. Moore e Denis Zuev, "Internet traffic classification using bayesian analysis techniques," SIGMETRICS '05: Proceedings of the 2005 ACM SIGMETRICS international conference on Measurement and modeling of computer systems, pp. 50-60, 2005.

[57] Kuai Xu, Zhi-Li Zhang e Supratik Bhattacharyya, "Reducing unwanted traffic in a backbone network," SRUTI'05: Proceedings of the Steps to Reducing Unwanted Traffic on the Internet on Steps to Reducing Unwanted Traffic on the Internet Workshop, p. 2, 2005.

[58] Varun Chandola, Arindam Banerjee e Vipin Kumar, "Anomaly detection: A survey," *ACM Comput. Surv.*, vol. 41, no. 3, pp. 1-58, 2009.

[59] Zengyou He, Xiaofei Xu e Shengchun Deng, "An Optimization Model for Outlier Detection in Categorical Data," *Advances in Intelligent Computing, 400-409*, vol. 3644, pp. 400-409, 2005.

[60] Shin Ando, "Clustering Needles in a Haystack: An Information Theoretic Analysis of Minority and Outlier Detection," *Data Mining, 2007. ICDM 2007. Seventh IEEE International Conference on*, pp. 13-22, 28-31, Outubro 2007.

[61] Sidney C. de Lucena e Alex Soares de Moura., "Análise dos Estimadores EWMA e Holt-Winters para Detecção de Anomalias em Tráfego IP a partir de Medidas de

Entropia," *CSBC 2009. WPerformance - VIII Workshop em Desempenho de Sistemas Computacionais e de Comunicação, 2009*, pp. 2177-2192, 2009.

[62] Guthemberg da Silva Silvestre, "Uma análise extensiva do tráfego de aplicações," Universidade Federal de Pernambuco, Dissertação de Mestrado 2007.

[63] Yan Ruo-Yu e Zheng Qing-Hua, "Multi-scale Entropy Based Traffic Analysis and Anomaly Detection," *Intelligent Systems Design and Applications, 2008. ISDA '08. Eighth International Conference on*, pp. 151-157, Novembro 2008.

[64] Libpcap, Libpcap, 2008.

[65] Antonis Papadogiannakis, Demetres Antoniades, Michalis Polychronakis, and Evangelos P. Markatos, "Improving the Performance of Passive Network Monitoring Applications using Locality Buffering," *MASCOTS '07: Proceedings of the 2007 15th International Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems*, pp. 151-157, Outubro 2007.

[66] J. Quittek, T. Zseby, B. Claise e S. Zander, Requirements for IP Flow Information Export (IPFIX), 2004.

[67] B. Claise, Cisco Systems NetFlow Services Export Version 9, 2004.

[68] Guangxing Zhang et al., "Accurate Online Traffic Classification with Multi-Phases Identification Methodology," *Consumer Communications and Networking Conference*, pp. 141-146, Janeiro 2008.

[69] Fang Hao, M. Kodialam, T.V. Lakshman e Hui Zhang, "Fast payload-based flow estimation for traffic monitoring and network security," *Architecture for networking and communications systems, 2005.* , pp. 211-220, Outubro 2005.

[70] Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest e Clifford Stein, "Chapter 11 - Hash Tables," in *Introduction to Algorithms, Second Edition.*: MIT Press, 2001.

[71] Internet Assigned Numbers Authority IANA. (2009) Top-Level Domains List. "http://data.iana.org/TLD/tlds-alpha-by-domain.txt". Acesso em 12/07/2009

[72] Microsoft. (2008) SRV Resource Records. http://technet.microsoft.com/pt-br/library/cc961719(en-us).aspx. Acesso em (23/05/2009)

[73] Cricket Liu e Paul Albitz, *DNS and BIND (5th Edition)*.: O'Reilly Media, Inc., 2006.

[74] David Dagon, "Botnet Detection and Response," in *OARC Workshop*, 2005.

Livros Grátis

(http://www.livrosgratis.com.br)

Milhares de Livros para Download:

Baixar livros de Literatura
Baixar livros de Literatura de Cordel
Baixar livros de Literatura Infantil
Baixar livros de Matemática
Baixar livros de Medicina
Baixar livros de Medicina Veterinária
Baixar livros de Meio Ambiente
Baixar livros de Meteorologia
Baixar Monografias e TCC
Baixar livros Multidisciplinar
Baixar livros de Música
Baixar livros de Psicologia
Baixar livros de Química
Baixar livros de Saúde Coletiva
Baixar livros de Serviço Social
Baixar livros de Sociologia
Baixar livros de Teologia
Baixar livros de Trabalho
Baixar livros de Turismo