Identificação e Caracterização de Comportamentos Suspeitos Através da Análise do Tráfego DNS

Kaio R. S. Barbosa[1], **Eduardo Souto**[1], **Eduardo Feitosa**[1], **Gilbert B. Martins**[1]

[1]Instituto de Computação - Universidade Federal do Amazonas (UFAM)
Av. Rodrigo Otávio Jordão Ramos, 3000, Coroado.
CEP 69077-000, Manaus-AM, Brasil

{kaiorafael,esouto,efeitosa,gilbert.martins}@icomp.ufam.edu.br

Abstract. *The Domain Name System (DNS) provides mechanisms for translating domain names into IP address. This service is used by both legitimate users and suspicious applications which may request mail servers' address before sending spam. This paper presents a methodology based on graph theory that distinguishes between legitimate and malicious traffic queries patterns. Name resolutions are modeled in a graph that illustrates the communication patterns between hosts and how the queries were held. To validate the proposal, the .br DNS domain traffic is investigated. The results show a 35% reduction of the hosts to be analyzed and the presence of suspicious behavior.*

Resumo. *O Sistema de Nomes de Domínios (DNS) fornece mecanismos para traduzir os nomes de domínios em endereços IP. Este serviço é usado tanto por usuários legítimos quanto aplicações suspeitas que podem solicitar endereços dos servidores de email antes de enviar spam. Este trabalho apresenta uma metodologia que utiliza teoria dos grafos para distinguir padrões entre consultas legítimas e maliciosas no tráfego. As resoluções de nomes são modeladas em um grafo que ilustra o padrão de comunicação entre hosts e como as consultas foram realizadas. Para validação da proposta, o tráfego DNS do domínio .br é investigado. Os resultados mostram uma redução de 35% do total de hosts a serem analisados e a presença de comportamentos suspeitos.*

1. Introdução

O Sistema de Nomes de Domínios (DNS) [Mockapetris 1987] traduz nomes de domínios em endereços IP permitindo que aplicações e usuários tenham acesso aos diversos serviços de rede como sistemas de comércio eletrônico, navegação de sites e envio de emails. Entretanto, devido ao acesso livre e distribuído do protocolo DNS, aplicações maliciosas também podem fazer consultas de nomes de domínios para realizar ataques, tais como negação de serviço, propagação de spam em massa e distribuição de *malwares*. Em todos os casos, o tráfego DNS é consultado inicialmente para obter informações sobre o serviço desejado.

A identificação e distinção de comportamentos benignos e maliciosos no tráfego DNS ainda é um problema em aberto, especialmente quando o tráfego é analisado em servidores de Domínio de Primeiro Nível, também conhecidos como *Top-Level Domains - TLD* e servidores Raiz (*Root Servers*) [Castro et al. 2008]. As possíveis razões são devidas ao volume de tráfego a ser processado ou a questões legais de acesso

 ©2014 SBC — Soc. Bras. de Computação

[Antonakakis et al. 2010]. Além disso, a distinção de comportamentos é agravada quando aplicações maliciosas mudam de comportamento para dificultar a sua detecção. Um típico exemplo de tal situação é o *worm Conficker* que acessa uma lista diária de novos domínios gerados dinamicamente por algoritmo [Shin e Gu 2010].

Embora algumas aplicações maliciosas consigam evadir os sistemas de detecção, ainda é possível detectar comportamento anômalo no tráfego DNS devido ao ciclo de vida das máquinas infectadas (*bots*), pois em algum momento, os *bots* utilizam o tráfego DNS para coletar informações antes do ataque. Desta forma, é possível assumir que hosts maliciosos apresentem padrões de comportamentos similares na rede. Por exemplo, antes de enviar um email, é necessário que o atacante realize uma consulta DNS para obter o endereço do servidor que recebe esse tipo de mensagem. Um abuso nesse tipo de consulta pode denotar o comportamento de *SpamBot* [Ishibashi et al. 2005]. Da mesma forma, um ataque de reconhecimento de rede, o tráfego DNS é consultado para identificar hosts que são válidos e ativos [Barbosa e Souto 2009].

Motivado pelos problemas citados acima, este trabalho apresenta uma metodologia que utiliza Teoria dos Grafos para identificar padrões de comportamentos suspeitos em tráfego TLD. Elementos da consulta DNS como IP de origem, nome de domínio e registro de recurso (RR - *resource record*) são modelados em um grafo direcionado que corresponde à consulta partindo do endereço IP de origem para o nome de domínio, e o registro de recurso denota o objetivo da consulta realizada. A metodologia proposta permite aos operadores de rede entender a relação de comunicação entre hosts e facilmente indicar aqueles hosts que possuem maior relevância para investigação. Tal abordagem minimiza não só a quantidade de hosts a ser analisada bem como o tempo necessário para a análise.

Para avaliar a metodologia proposta, este trabalho investiga as consultas recebidas por servidores de primeiro nível (TLDs) do domínio `.br`, disponíveis no projeto DITL (*Day in The Life of the Internet*)[DITL 2014]. Consultas DNS em TLDs fornecem uma visão em larga escala de como o tráfego DNS de uma região é utilizado. Por exemplo, caso um host envie spam em massa para diferentes regiões geográficas no globo, os destinatários deste email solicitarão do servidor de domínio do país de origem (TLD) mais informações sobre o remetente. Desta forma, ao observar consultas DNS em servidores de domínios de primeiro nível, é possível entender como um domínio é requisitado por hosts na Internet e por hosts dentro da zona de domínio.

Diferente de outros trabalhos, a metodologia proposta aqui utiliza apenas as solicitações DNS como fonte de análise. Além disso, os resultados mostram que a metodologia proposta reduz em média 35% dos hosts a serem analisados por operadores de redes. Devido à atividades supostamente suspeitas, clientes de banda larga - usuários domésticos - são os principais hosts consultados por servidores de email e servidores recursivos DNS (rDNS). Uma validação mostrou que 93% dos hosts identificados como relevantes estavam cadastrados em listas negras. Da mesma forma, é possível identificar um conjunto de características que podem ser utilizadas para geração de filtros e classificação automática de tráfego anômalo.

O restante desse trabalho está organizado da seguinte forma: a Seção 2 apresenta os trabalhos relacionados à identificação de comportamentos suspeitos através do

 ©2014 SBC — Soc. Bras. de Computação

tráfego DNS; a Seção 3 demonstra como a Teoria dos Grafos é utilizada no processo de detecção de anomalias; a Seção 4 apresenta os resultados obtidos, descrevendo o padrão de comunicações e evidências de *bots* para o envio de spam; a Seção 5 apresenta as conclusões sobre os resultados e apresenta novos direcionamentos e trabalhos futuros.

2. Trabalhos Relacionados

A análise passiva do tráfego DNS de servidores raiz (*root servers*), servidores de Domínios de Primeiro Nível (chamados de *Top-Level Domains*, ou TLDs) como `.com`, `.net` e `.gov`, e de servidores responsáveis pelos domínios de código de país (*country code TLDs*, ou ccTLDs) permite entender como a resolução de nomes é utilizada globalmente.

Muitos trabalhos têm como objetivo entender e identificar as principais características do tráfego de domínio de primeiro nível, incluindo frequência de registros de recurso, consultas mal formadas e domínios inválidos. Por exemplo, para entender como o tráfego DNS é utilizado por servidores raiz, Castro et al. (2008) analisam o tráfego DNS de servidores raiz coletado em Janeiro de 2006, Janeiro de 2007 e Março 2008 através do projeto DITL (*Day in the Life of the Internet*)[DITL 2014]. Os resultados mostram como os registros de recursos são distribuídos entre as consultas e que existe uma grande quantidade de consultas inválidas que não deveriam alcançar esses servidores, como consultas com TLDs inválidos, consultas recursivas e consultas onde o nome do domínio é um endereço IP (`A-to-A`).

Outros trabalhos direcionam essa análise para servidores de domínios de código de país. Em 2009, Barbosa e Souto (2009) analisaram o tráfego do domínio `.br` coletado em dois dias do mês de março de 2008. Entre os resultados, os autores observaram que as consultas do tipo `PTR` , utilizadas para informar o nome de um domínio a partir de um endereço IP, representam 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. Devido ao volume de tráfego `PTR` em relação ao tipo `A` , os autores sugeriram que tal comportamento era resultado de atividades suspeitas no tráfego, pois outros trabalhos que investigam o tráfego TLD apontam que o registro `A` é o mais frequente [Castro et al. 2008, Yuchi et al. 2009]. Por esse motivo, este trabalho segue com uma análise mais detalhada das consultas destinadas ao domínio `.br` observando agora tráfegos coletados no projeto DITL em 2008, 2009 e 2010 [1].

Antonakakis et al. (2010) apresentam um sistema de reputação dinâmica para nomes de domínios novos ou desconhecidos no ccTLD do Canadá (`.ca`). O sistema, denominado `Notos`, analisa o comportamento do tráfego DNS e atribui uma pontuação de acordo com as atividades relacionadas com o domínio investigado. O comportamento histórico do DNS é obtido a partir de bases legítimas e maliciosas. O comportamento legítimo é coletado em servidores recursivos DNS, enquanto o tráfego malicioso é obtido através de sensores de rede como *honeypots*, *spam-traps* e *sandboxes*. Para facilitar a análise do tráfego, os autores agrupam domínios de comportamentos semelhantes observando características léxicas do nome de domínio (e.g.: frequência de caracteres e tamanho do nome) e de rede (e.g.: número AS e total de endereços IPs associados ao

[1] O Brasil participou do projeto DITL até o ano de 2010.

 ©2014 SBC — Soc. Bras. de Computação

domínio). A partir dessas características, nomes de domínios que apresentem os comportamentos conhecidos podem ser classificados como benignos ou suspeitos. O trabalho de Choi e Lee (2012) também utiliza o tráfego DNS para agrupar comportamentos semelhantes. Os autores observam consultas repentinas, quantidade de consultas únicas, domínio requisitado por múltiplos IPs de origem em um determinado intervalo de tempo e tentam identificar nesses padrões de consultas redes maliciosas (*botnets*).

Ramachandran et al. (2006) utilizaram análise de grafo para identificar consultas às listas negras originadas por membros de *botnets*. Os autores observaram a relação entre o número de consultas enviadas e recebidas por essas listas. A partir dessa relação, os autores determinaram limiares que permitiram identificar comportamento suspeito no tráfego DNS. Ramachandran et al. validaram a proposta utilizando uma base de *spamtrap*, a qual possuía endereços de *bots* conhecidos. Tais endereços foram observados consultando listas negras para identificar se outros *bots* da mesma rede, ou o próprio endereço, estavam cadastrados nessas bases de dados.

De maneira geral, este trabalho tem como objetivo entender o padrão de comunicação entre os hosts na rede através do protocolo DNS e como esse tráfego é utilizado. Para entender esse padrão, os registros de recursos são utilizados como filtro para identificar e extrair comportamentos de hosts relevantes para análise. Por exemplo, o registro `MX`, que indica uma lista de servidores que devem receber emails de um domínio, é frequentemente utilizado para encontrar servidores de email abertos e suscetíveis aos ataques de spam em massa [Ishibashi et al. 2005]. Em alguns casos, o tipo `A` também pode ser utilizado para detectar domínios de spam [Jiang et al. 2010]. O registro reverso (`PTR`) é útil para identificar endereços IPs válidos, pois cada IP acessível na Internet deve possuir um nome reverso [Barr 1996]. Em ataques de reconhecimento de rede, o registro `PTR` é mais utilizado [Kumagai et al. 2010]. Ataques de dicionário contra o serviço `SSH` podem ser identificados através da distribuição de frequência do registro reverso [Shibata et al. 2012]. Por esses motivos, os tipos de registros de recursos são essenciais para detectar anomalias que utilizam o tráfego DNS.

3. Metodologia

A metodologia proposta consiste em um modelo de representação do tráfego de rede baseada em grafos direcionados que modelam os relacionamentos entre hosts e domínios consultados. As etapas do processo de identificação de consultas maliciosas a partir da análise do tráfego DNS são:

1. Construção do grafo original, onde o tráfego DNS é modelado em grafo direcionado. Os vértices são formados por hosts e nomes de domínios e as arestas, as comunicações entre os vértices.
2. Transformação do grafo, tem objetivo reforçar as conexões do grafo para encontrar padrões de comunicações que não foram modelados inicialmente.
3. Redução do grafo, onde os componentes conexos irrelevantes do grafo transformado são eliminados.
4. Classificação das consultas, onde um conjunto de métricas, definidas para descrever as propriedades estruturais do grafo, é usado para classificar os nós e identificar possíveis comportamentos maliciosos nos hosts associados.

 ©2014 SBC — Soc. Bras. de Computação

3.1. Construção do Grafo Original

As consultas DNS são modeladas através de um grafo $G = (V, E)$ direcionado, onde V e E são os conjuntos de vértices e arestas de G, respectivamente. Seja A o conjunto de endereços IP de origem e D o conjunto de nomes de domínio. Então, seja $V = A \cup D$, tal que $a \in A$, representa o endereço IP origem da consulta; e $d \in D$ representa o nome do domínio da consulta. Uma aresta $e \in E$, onde $e = (a, d)$, denota a consulta partindo do endereço IP de origem a para o domínio da consulta d. Uma aresta possui o atributo t, tal que $(a, d).t$ denota o tipo de registro de recurso associado à consulta de a para d. Seja $v \in V$, tal que $deg^+(v)$ denota o grau de saída de um vértice e $deg^-(v)$ denota o grau de entrada. O grau do nó $deg(v)$ é denotado pela soma de todos os graus de entrada e saída do vértice. Finalmente, a função $f_{ip}(v)$ retorna o endereço IP de $v \in V$.

Para ilustrar como as consultas DNS são modeladas em um grafo direcionado, considere o exemplo na Figura 1. Por questões de privacidade, endereços privados serão utilizados. A Figura 1 apresenta um conjunto de *Consultas DNS* e o grafo resultante da modelagem, denominado como *Grafo Original*. Por simplicidade, os endereços IPs `192.168.0.1`, `192.168.1.1` e `192.168.3.1` são denotados como `A1`, `A2` e `A3`, e os domínios da consulta `google.com.br`, `example.com`, `1.3.168.192.in-addr.arpa` e `sbc.org.br` são denotados como `D1`, `D2`, `D3`, `D4`, respectivamente. Na primeira linha das consultas DNS, é possível observar o endereço IP `A1` utilizando o registro do tipo `MX` para encontrar o endereço do servidor de email que responde pelo domínio `google.com.br`. Tal consulta é ilustrada pela aresta $(A1, D1)$ no *Grafo Original*. A segunda consulta realizada por `A1` é demonstrada na terceira linha. Tal exemplo denota `A1` buscando pelo endereço reverso (`PTR`) de `1.3.168.192.in-addr.arpa`. A representação dessa consulta é demonstrada pela aresta $(A1, D3)$.

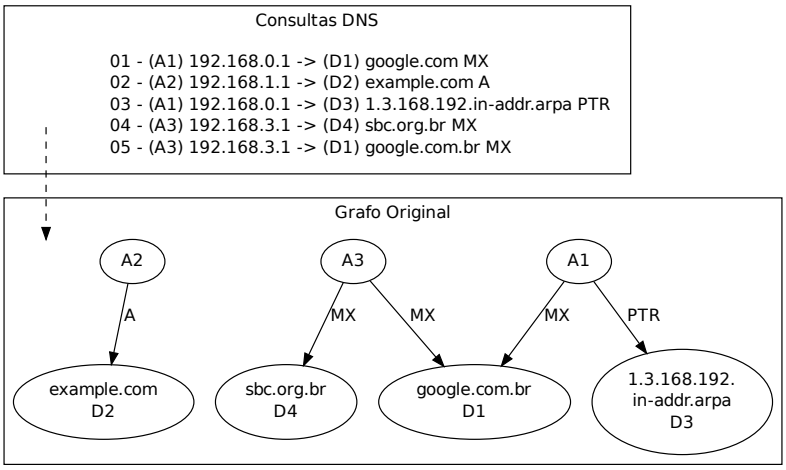

Figura 1. Exemplos de consultas DNS modeladas em Grafo.

3.2. Transformação do Grafo Original

Após a modelagem inicial do grafo, é necessário identificar comportamentos que não foram modelados no *Grafo Original*. Considere o exemplo da consulta na terceira linha em *Consultas DNS*. É possível observar que o nome de domínio (`D3`) foi consultado a partir do registro `PTR` , por isso para identificar se $D3$ também é um endereço IP de origem, isto é, se o host também enviou solicitações, o nome de domínio é convertido do

 ©2014 SBC — Soc. Bras. de Computação

formato reverso `d.c.b.a.in-addr.arpa` para o endereço IP original `a.b.c.d`. A transformação do grafo tem como objetivo reforçar as conexões dos vértices para representar o padrão de comunicação entre hosts mais próximo da realidade.

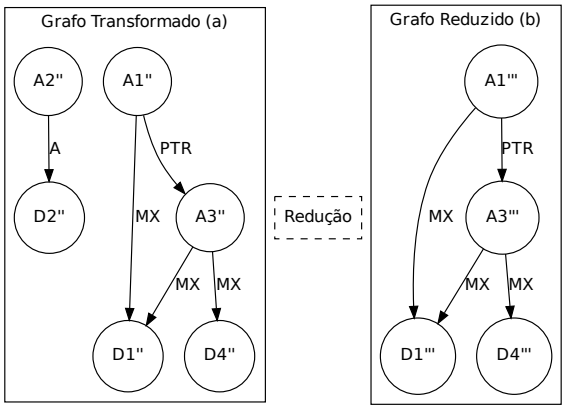

Figura 2. Grafo transformado (a) da Figura 1 e grafo reduzido (b).

Considerando o exemplo da Figura 1, o domínio `D3` equivale ao endereço IP `A3` no formato reverso. Portanto, o grau de entrada de $A3$ no *Grafo Original* resulta em $deg^{-}(A3) = 0$, e o grau de saída $deg^{+}(A3) = 2$. Por outro lado, o grau de $A3$ no *Grafo Transformado* da Figura 2a demonstra o resultado $deg^{-}(A3) = 1$ e $deg^{+}(A3) = 2$. Formalmente, o processo de transformação do grafo é uma função definida pelas seguintes propriedades:

Definição 1: Seja $G = (V, E)$ o *Grafo Original*. O conjunto R é uma coleção de pares ordenados (a_i, a_k), com a_i e $a_k \in V$, gerado pela função $f_1 : V \rightarrow R$ tal que, para todo par ordenado $(a_i, a_k) \in R$ existe um par $(a_i, d_j) \in E$ que foi derivado de uma consulta à um nome de domínio d_j do tipo `PTR`, onde $f_{ip}(a_k) = f_{ip}(d_j)$.

Definição 2: Seja $G = (V, E)$ o *Grafo Original*. Seja R o conjunto gerado de acordo com a Definição 1. O conjunto S é gerado pela função $f_2 : (D, R) \rightarrow S$ tal que, $s \in V$ e para todo $s \in S$ existe um par $(a_i, a_k) \in R$ onde $f_{ip}(s) = f_{ip}(a_k)$.

Definição 3: Seja $G = (V, E)$ o *Grafo Original* e $G' = (V', E')$. Seja R o conjunto gerado de acordo com a Definição 1. Seja S o conjunto gerado de acordo com a Definição 2. Então $G' = (V', E')$ é definido pela função geradora $f_3 : (G, R, S) \rightarrow G'$, tal que $V' = V - S$, $E' \subset (E \cup R)$ e, para todo $(a', d') \in E'$ e todo o $s \in S$, não existe $f_{ip}(d') = f_{ip}(s)$.

3.3. Redução do Grafo Transformado

Finalmente, após o processo de transformação do grafo, é necessário excluir consultas que não são representativas para análise, isto é, solicitações onde o endereço IP de origem comunica-se com o nome de domínio uma única vez, conforme aresta (A2, D2) do grafo transformado (Figura 2a). A Figura 2b ilustra o resultado da redução do grafo. Formalmente, se a_k e d_k constituem um componente conexo de G', então tais vértices são eliminados quando $deg(a_k) = 1$ e $deg(d_k) = 1$. Assim:

Definição 4: Seja $G' = (V', E')$ um grafo gerado de acordo com a Definição 3. O *Grafo Transformado* $G'' = (V'', E'')$ é gerado pela função $f_4 : G' \rightarrow G''$, onde $V'' = V'$

 ©2014 SBC — Soc. Bras. de Computação

e $E'' \subset E'$ tal que, para todo o par ordenado $(a'', d'') \in E''$, tem-se $deg(a'') > 1$ ou $deg(d'') > 1$.

Definição 5: Seja $G'' = (V'', E'')$ um grafo gerado de acordo com a Definição 4. O *Grafo Reduzido* $G''' = (V''', E''')$ é gerado pela função $f_5 : G'' \rightarrow G'''$, onde $E''' = E''$ e $V''' \subset V''$ tal que, para todo $v''' \in V'''$, tem-se $deg(v''') > 0$.

3.4. Classificação dos Comportamentos Relevantes

O entendimento e classificação de comportamentos através do tráfego DNS são obtidos pela relação entre hosts e como as consultas DNS foram realizadas. Neste trabalho a relação entre hosts é definida pela aresta de pares ordenados e os registros de recursos mostram como os hosts se comunicam. Formalmente, relações de hosts são consideradas relevantes para análise quando o host possui: i) alto grau de entrada do nó - são observados principalmente nós que possuem alta incidência de consultas dos registros de recurso do tipo `PTR` e `MX`, com exceção dos nós com alta incidência natural do tipo `A` como consultas ao `google.com.br` e `facebook.com.br`, pois não representam necessariamente comportamentos suspeitos; ii) alto grau de saída de nó - são observados nós que realizam grandes quantidades de consultas empregando principalmente os registros `PTR`, `MX` e `NS`[2]. Tais tipos de registros de recursos são comumente usados por hosts infectados (*bots*) em diversas atividades maliciosas como envio de mensagens de spam e ataques de reconhecimento. A relação entre hosts também é observada pela iii) razão entre o grau saída e entrada do vértice - a partir dessa relação é possível identificar padrões de consultas semelhantes no tráfego. Por exemplo, servidores de email legítimos são observados no tráfego enviando e recebendo consultas, no entanto, hosts acessando a Internet, geralmente, apenas enviam consultas [Ramachandran et al. 2006]

O entendimento de como as resoluções foram realizadas é obtida pela iv) frequência relativa dos registros de consultas enviadas; e a v) frequência relativa dos registros de consultas recebidas. Tais frequências permitem agrupar hosts que abusam dos registros (e.g: `PTR` e `MX`) independente da quantidade de consultas enviadas. Portanto, a partir desse conjunto de métricas é possível estabelecer uma correlação para definir classes de consultas.

3.5. Classes de Consultas

As Classes de Consultas representam a agregação de hosts relevantes identificados a partir das métricas aplicadas aos vértices. Formalmente, hosts são agrupados pela distribuição estatística de frequência das características individuais. Tal abordagem permite que diferentes comportamentos sejam correlacionados para entendimento do padrão de consultas, ou seja, é possível observar o host em momentos distintos no tráfego DNS. Por exemplo, a razão entre o número de consultas enviadas pelo número de consultas recebidas contribui na análise do tráfego quando a mesma é correlacionada com a frequência de registros de recurso enviados ou recebidos. Por esses motivos, os registros de recursos são fundamentais no entendimento do tráfego.

As classes de consultas são correlacionadas para demonstrar como um host utiliza o tráfego DNS para comunicação. A Tabela 1 apresenta a relação entre as métricas de grafo. Para cada intervalo da classe de consultas a ser investigado, uma métrica é definida

[2]O registro `NS` é utilizado para encontrar o servidor de nomes que responde pela zona de domínio.

 ©2014 SBC — Soc. Bras. de Computação

Tabela 1. Relação entre as métricas de grafo.

Classe Primária Q	**X**	**Y**	**Z**	**W**
Grau de Entrada (deg^-)	deg^+	Razão	Freq. RR Enviados	Freq. RR Recebidos
Grau de Saída (deg^+)	Razão	deg^-	Freq. RR Enviados	Freq. RR Recebidos
Razão Saída / Entrada	deg^-	deg^+	Freq. RR Enviados	Freq. RR Recebidos
Frequência Registros Recursos Enviados	deg^-	deg^+	Razão	Freq. RR Recebidos
Frequência Registros Recursos Recebidos	deg^-	deg^+	Razão	Freq. RR Enviados

como primária e comparada com as demais métricas do grafo. Por exemplo, seja $Q(deg^-)$ a classe primária, então um vértice é analisado com base no grau de saída (X), razão (Y) e frequência relativa dos registros de recursos enviados (Z) e recebidos (W). Formalmente, um comportamento de um vértice é definido por $Q_{(I(v'''))} = [X, Y, Z, W]$.

4. Resultados

Nesta Seção são demonstrados os resultados obtidos durante experimentos. Primeiramente, a base de dados utilizada na metodologia é descrita. Em seguida, o padrão de comunicação entre os hosts é demonstrado a partir do grau de saída e grau de entrada dos vértices. A classe primária do grau de entrada é utilizada para ilustrar como os hosts relevantes são observados no tráfego. Finalmente, os hosts relevantes são passivamente analisados para demonstrar a validade da metodologia proposta.

Para validar a metodologia proposta foi utilizado tráfego DNS real coletado durante o projeto DITL [DITL 2014], uma cortesia da DNS-OARC (Centro de Pesquisa, Operações e Análise DNS). O tráfego obtido é composto por 15 servidores DNS autoritativos que respondem pelo domínio `.br` , sendo cinco em 2008, quatro em 2009 e seis em 2010. Todos os servidores iniciam a coleta em 00:00:00 UTC e terminam em 23:59:59:9999. A Tabela 2 sumariza os valores na base de dados: dia de coleta; servidores de nome, total de pacotes (consultas e respostas) e o tamanho da base no formato compactado `gzip`.

Tabela 2. Base de dados DITL.

	DITL 2008	DITL 2009	DITL 2010
Dias de Coleta	[18-19]/03	[30-31]/03, 01/04	[14-15]/04
Servidores DNS	{a-e}.dns.br	{a,b,e,f}.dns.br	{a-f}.dns.br
Total de Pacotes	5.3 bilhões	6.4 bilhões	6.9 bilhões
Tamanho da base	229GB	236GB	282GB

Por questões de segurança e privacidade, a análise do tráfego DNS do `.br` deve ser realizada exclusivamente dentro da infraestrutura fornecida pela DNS-OARC. Além disso, tal infraestrutura é compartilhada entre outros pesquisadores que também investigam o tráfego DNS disponível. Por esse motivo, este trabalho investiga o tráfego DNS do servidor `a.dns.br` durante o período 00:00 até 00:59 dos dias 18/03, 30/03 e 14/04 de 2008, 2009 e 2010, respectivamente.

4.1. Características dos Hosts Relevantes

A Tabela 3 apresenta os valores observados durante aplicação da metodologia proposta. O *Total de Consultas* representa todas as consultas observadas durante T, tal que $T = 1h$.

 ©2014 SBC — Soc. Bras. de Computação

O *Total de Hosts Únicos* demonstra o número de hosts na base antes da identificação do *Total de Hosts Relevantes*. Por exemplo, em 2008 o total de hosts relevantes é 47% menor que o total de hosts únicos. Em 2009 e 2010 essa redução representa 49% e 10%, respectivamente. Isto é, através da metodologia proposta foi possível reduzir em média 35% do total de hosts a serem analisados. Finalmente, a distribuição de frequência dos registros de recurso utilizados pelos hosts relevantes mostra que o registro `PTR` é o mais frequente em comparação aos registros `A` e `MX`.

Tabela 3. Resumo dos hosts relevantes encontrados para análise.

	DITL 2008	DITL 2009	DITL 2010
Total de Consultas	9.985.841	23.190.993	25.193.658
Total de Hosts Únicos	263.036	351.353	342.045
Total de Hosts Relevantes	138.558	178.519	306.124
Dist. Freq. - A|PTR|MX	32% | 53% | 12%	20% | 74% | 4%	39% | 43% | 16%

Para entender o padrão de comunicação entre os hosts identificados como relevantes, a Figura 3 apresenta o grau de saída e o grau de entrada dos vértices analisados em 2008, 2009 e 2010. Por razões de espaço, 20 mil hosts de cada ano foram escolhidos aleatoriamente e plotados. É possível destacar que muitos hosts são observados apenas enviando ou recebendo solicitações, visto que, os valores assumidos estão concentrados nos eixos X ou Y. De maneira geral, em 2008 e 2009, o grau de saída dos hosts está abaixo de 5 mil consultas enviadas, enquanto em 2010, esse padrão é divergente. Os valores identificados mostram que 22 hosts enviaram mais de 60 mil requisições e, por isso, foram investigados. Tais hosts estão distribuídos ao longo de sistemas autônomos de provedores de serviço de banda larga e provedores de Internet. O número AS foi obtido através do mapeamento `IP-to-ASN` disponível no projeto *Team CYMRU*.[3]

Para o grau de entrada, a maioria dos hosts recebeu menos de 600 consultas durante T, embora em 2010 seja possível observar hosts divergindo desse padrão. Uma análise superficial de tais hosts mostra que os nomes de domínios são consultados a partir seu nome reverso (`host.in-addr.arpa`), e que esses endereços correspondem a 78% de endereços de banda larga.

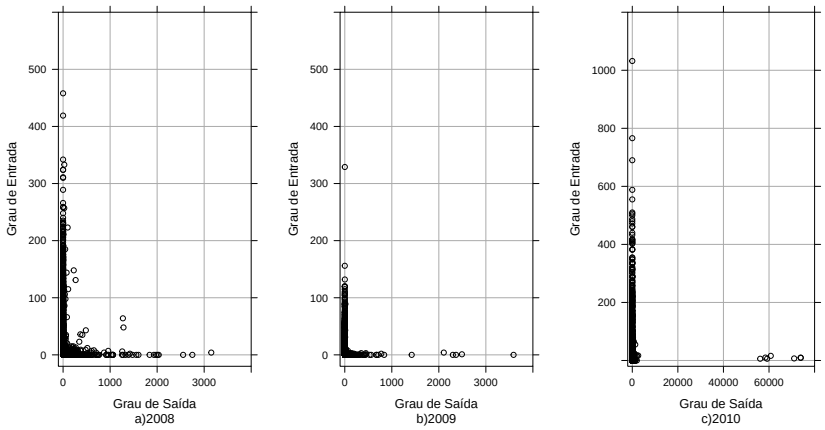

Figura 3. Visão geral das consultas enviadas ou recebidas na base de dados DITL 2008, 2009 e 2010.

[3]http://www.team-cymru.org/Services/ip-to-asn.html

 ©2014 SBC — Soc. Bras. de Computação

Para ilustrar o comportamento dos hosts que enviaram mais de 60 mil consultas, a Tabela 4 apresenta o resumo das principais características dos 5 primeiros hosts desse grupo. Nessa tabela são apresentados o grau de entrada, grau de saída, razão, frequência dos registros recebidos e frequência dos registros enviados, respectivamente. A frequência dos registros de recursos de entrada denotam 100% de consultas do tipo `PTR`, com exceção do host de número 4, o qual recebeu consultas `PTR` (92.4%) e do tipo $\star$ (7.6%) [4]. As frequências dos registros de saída mostram os valores dos registros do tipo `A`, `PTR` e `MX`, entre os quais o tipo `A` é o mais frequente.

Tabela 4. Hosts relevantes em 2010 que enviaram mais de 60 mil consultas DNS.

Host	deg^-	deg^+	Razão	Freq. RR Entrada	Freq. RR Saída
1	50	204.616	4092.32	100%	75.2%| 14.7%| 14.4%
2	45	177.259	3930.08	100%	73.6%| 10.8%| 15.5%
3	37	130.205	3519.05	100%	70.3%| 9.4%| 19.8%
4	13	105.763	8135.61	92.4% | 7.6%	74.4%| 6.6%| 18.8%
5	37	97.858	2644.81	100%	81.1%| 6.9%| 11.8%

4.2. Classe Primária - Grau de Entrada

Para ilustrar como os hosts são observados nas classes de consultas, a classe primária do grau de entrada é demonstrada a seguir. Os 100 primeiros hosts que tiveram maior incidência de consultas foram identificados e atribuídos em intervalos distintos das classes de comportamentos. A Figura 4 apresenta a frequência total dos hosts em cada classe. Em 2008 e 2009, praticamente 100% dos hosts estavam localizados abaixo de 1000 consultas. Entretanto, em 2010, esse valor representa 60%. Devido à base de 2010 ser a mais recente, os valores encontrados serão descritos a seguir.

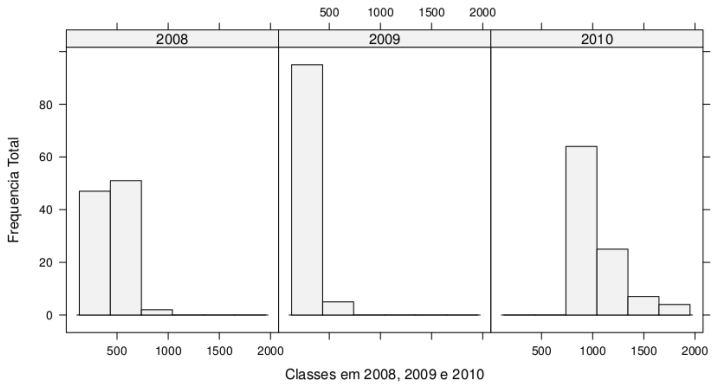

Figura 4. Intervalo de Classes para cada ano da base de dados.

Em 2010, 60% dos hosts que possuem maior incidência de consultas estão dentro do intervalo entre 500 e 1000 consultas. Nesse intervalo de classe, 83% dos hosts possuem o nome reverso apontando para endereços IPs e sistemas autônomos de banda larga. Os demais endereços são nomes no formato reverso que possuem o domínio inexistente (`NXDomain`). Do conjunto de hosts identificados nesse intervalo, vale destacar dois grupos; i) o conjunto de hosts que apenas recebem consultas; e o ii) conjunto de hosts que

[4]Consultas do tipo $\star$ são utilizadas para obter informações de todos os registros disponíveis do nome de domínio.

 ©2014 SBC — Soc. Bras. de Computação

enviam e recebem consultas. O primeiro grupo representa 70% do total de hosts analisados. Tais hosts são consultados através dos registros de recurso `PTR`, $\star$ e `CNAME`[5]. O grau de saída desse grupo denota valores entre 0 e 1 consulta, os quais podem resultar na Razão igual a zero ou próxima de zero, respectivamente.

O segundo grupo representa 30% do intervalo citado e a frequência dos registros de recursos de saída denota comportamento suspeito. Por exemplo, os hosts identificados nesse grupo abusam do registro de recurso do tipo `MX`. Em média, esse registro representa 79% do total de consultas enviadas, enquanto os registros do tipo `A` e `PTR` representam 20% e 1%, respectivamente. Além disso, em comparação com o grau de saída (deg^+) dos hosts da Seção 4.1, os hosts do grupo ii apresentam volume baixo de consultas enviadas, em média esses hosts enviaram 231 solicitações DNS, das quais o tipo `MX` era predominante. A seguir é realizado uma análise passiva dos hosts encontrados nos conjuntos i e ii.

4.2.1. Análise Passiva dos Hosts Relevantes

Para validar a relevância dos hosts identificados nas Classes de Consultas, a análise passiva pode ser utilizada como uma alternativa. Informações adicionais sobre o vértice são utilizadas para determinar se o comportamento é benigno ou malicioso. Por exemplo, através da análise léxica do nome de domínio, um endereço IP é identificado como banda larga caso o nome de domínio seja composto por palavras como `cpe`, `vivax`, `adsl`, `modem`, `cliente`, `dial-up`, `dyn`, `dynamic`, `dsl`, `brasiltelecom`, `gvt`, `velox`, `user.vivozap` e `virtua`. Diversos trabalhos demonstram que clientes de banda larga - usuários domésticos - são as principais vítimas de ataques [Dagon et al. 2007, Antonakakis et al. 2010, Choi e Lee 2012].

De maneira semelhante, através da análise léxica, um servidor de email ou servidor de nomes é identificado caso o nome de domínio apresente palavras como `mail`,`exchange`, `pop`,`pop3`, `imap` ou `ns`, `dns` e `nameserver`, respectivamente. Desta forma, este trabalho assume como comportamento suspeito um endereço de banda larga abusando dos registros `PTR` ou `MX`, ou ainda, quando múltiplos servidores de email, dentro do mesmo espaço de tempo, consultam por um endereço de banda larga.

Para ilustrar o comportamento citado, considere os hosts relevantes que apenas receberam consultas, conforme descrição na Seção 4.2. Do total de 70 hosts, 88.5% são endereços de banda larga e os demais denotam nome de domínio inexistente (`NXDomain`). Uma verificação superficial mostrou que 85.8% dos hosts estavam cadastrados em listas negras por não seguirem as definições da RFC 2142 (1997). Tal RFC prevê um conjunto de endereços de emails especiais que um domínio deve possuir para atender reclamações de abuso ou de entrega de email. Além disso, a RFC 2142 encoraja que servidores de email implementem pelo menos uma caixa de correio capaz de lidar com os problemas citados [Crocker 1997]. Isto é, essa RFC estipula diretrizes de como um serviço de email deve operar corretamente na Internet. No entanto, para que os hosts sejam reconhecidos como servidores de email, o registro reverso deveria apontar para um nome de domínio diferente do nome de provedores de banda larga, caso contrário a resolução reversa não estaria seguindo a RFC 1912 [Barr 1996].

Adicionalmente, vale destacar que esses hosts possuem padrões semelhantes de

[5]É usado para especificar que um nome de domínio usa o endereço IP de outro domínio.

 ©2014 SBC — Soc. Bras. de Computação

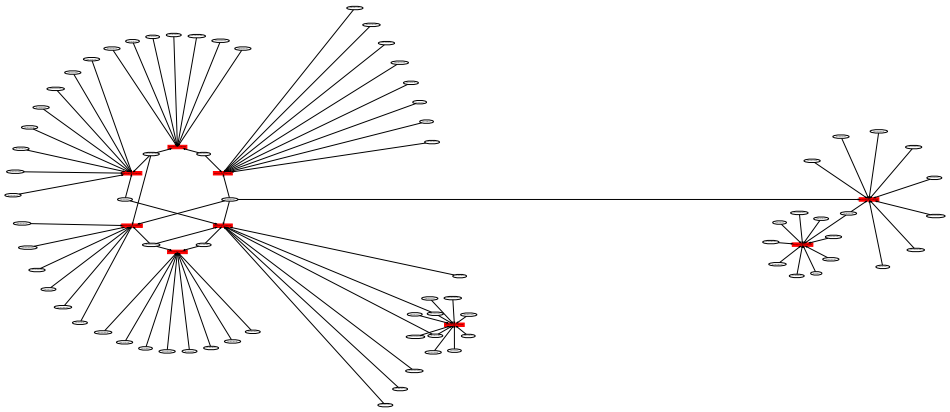

Figura 5. Padrão de consultas recebidas por *bots* para envio de spam. Quadrados vermelhos denotam nome da consulta, enquanto as elipses representam o IP de origem.

consultas recebidas. Primeiro, a maioria dos endereços IPs de origem que consultaram esses hosts são servidores de email e DNS. Segundo, o grau de entrada durante $\alpha = 10s$, tal que $\alpha \in T$, denota que esses hosts receberam em média 6 consultas a partir de endereços IPs distintos. Finalmente, é possível observar que um conjunto de hosts relevantes são consultados por um ou mais endereços IP de origem. Por questões de espaço, uma pequena parte desse padrão de comunicação é ilustrada na Figura 5. Nós vermelhos denotam os hosts frequentemente consultados e os nós em elipse representam os servidores de email ou de DNS. Diante desse contexto, é possível assumir que os hosts relevantes estão sendo consultados e validados após o envio em massa de spam.

```
01 - 201.b.c.x2z <domínio>.com.br MX
02 - 201.b.c.x2z ns1.<domínio>.com.br A
03 - 201.b.c.x2z itajuba.com.br MX
04 - 201.b.c.x2z <domínio>.com.br MX
05 - 201.b.c.x2z <domínio>.com.br MX
...
06 - 200.202.252.2 d.c.b.a.in-addr.arpa PTR
07 - 200.202.252.2 d.c.b.a.in-addr.arpa PTR
08 - 200.202.252.2 x2z.c.b.201.in-addr.arpa PTR
09 - 200.202.252.2 d.c.b.a.in-addr.arpa PTR
10 - 200.202.252.2 d.c.b.a.in-addr.arpa PTR
```

Figura 6. Exemplo de consultas realizadas por um *bot* para envio de spam.

Para comprovar o comportamento de envio de spam, a Figura 6 apresenta um exemplo dos hosts relevantes que enviaram e receberam consultas, conforme descrição na Seção 4.2. No exemplo é possível observar o endereço IP de origem, nome da consulta e o tipo de registro de recurso utilizado. As linhas 01-05 demonstram consultas do tipo `MX` e `A`. A linha 03 ilustra a consulta de serviço de correio eletrônico referente ao domínio `itajuba.com.br`. Em seguida, na linha 08, o endereço IP `x2z.c.b.201.in-addr.arpa` é solicitado pelo servidor de e-mail do domínio em questão (`200.202.252.2`). Tal comportamento pode ser identificado caso exista uma aresta entre o endereço IP do servidor de email e o endereço IP de origem inicial da consulta DNS. No entanto, para obter o endereço IP do servidor de email é necessário realizar uma nova consulta, a qual aumenta o tempo de análise dos hosts relevantes.

Além do padrão das consultas, outras características podem ser observadas nos

 ©2014 SBC — Soc. Bras. de Computação

hosts relevantes. Tais máquinas enviam consultas com tempo de vida (TTL) dentro do intervalo [118,125], bit de fragmentação do IP (DF) ausente e marcam o bit de recursividade (RD). Em uma breve análise em ambiente virtualizado, é possível constatar que máquinas com Windows XP não marcam o bit DF nas consultas DNS, enquanto sistemas baseados em Unix habilitam esse bit. Computadores funcionando com o Windows também são identificados no tráfego DNS através do TTL indicado [Wessels e Fomenkov 2003]. A identificação de hosts usando Windows é útil para detectar máquinas comprometidas, pois esse sistema operacional possui o maior número de infecções.

Consultas recursivas (bit RD) não deveriam alcançar servidores de raiz, pois tais tipos de resoluções não são atendidas por servidores de primeiro nível [Castro et al. 2008]. Consultas recursivas também são resultado de aplicações maliciosas que podem utilizar sistemas de resolução independentes para evadir dos sensores de rede. Portanto, é possível assumir que tais máquinas estão utilizando o tráfego DNS para propagação de spam. A validação desse comportamento suspeito pode ser obtida através de consultas às listas negras, onde 93% dos hosts identificados ainda são encontrados nas bases de dados.

5. Conclusão

Neste trabalho foi apresentada uma metodologia para a identificação e classificação de anomalias de rede a partir da correlação das consultas no tráfego DNS. As consultas DNS foram modeladas em um grafo direcionado e o registro de recurso PTR utilizado para reforçar as conexões, resultando em subgrafos mais densos. Durante os experimentos foi identificado que os hosts que tinham maior incidência de consultas eram clientes de banda larga e que através da metodologia proposta, o volume de tráfego a ser analisado era em média 35% menor. A principal vantagem da metodologia proposta é a sua fácil implementação, pois apenas as solicitações DNS são utilizadas para indiciar comportamentos suspeitos.

Como trabalhos futuros, outros registros do tráfego DNS podem ser utilizados para identificar comportamentos suspeitos. Por exemplo, o tamanho das consultas utilizando o registro EDNS é útil para identificar ataques de amplificação de resposta DNS. Os registros CNAME e TXT podem ser utilizados para detecção de tunelamento de tráfego através do protocolo DNS. Finalmente, algoritmos de aprendizagem de máquina seriam utilizados para detectar e classificar comportamentos suspeitos em tempo real a partir do emprego das características extraídas dos hosts relevantes.

Agradecimentos

Este trabalho agradece DNS-OARC e ao *Registro.br* pelo acesso ao tráfego. Este trabalho foi desenvolvido com o apoio do Governo do Estado do Amazonas por meio Fundação de Amparo à Pesquisa do Estado do Amazonas, com a concessão de bolsa de estudo.

Referências

Antonakakis, M., Perdisci, R., Dagon, D., Lee, W., e Feamster, N. (2010). Building a dynamic reputation system for dns. In *Proceedings of the 19th USENIX conference on Security*, USENIX Security'10, pages 18–18, Berkeley, CA, USA. USENIX Association.

Barbosa, K. R. S. e Souto, E. (2009). Análise passiva do tráfego dns da internet brasileira. In *IX Simpósio Brasileiro em Segurança da Informação e de Sistemas Computacionais - SBSeg 2009*, pages 203–216, Campinas.

 ©2014 SBC — Soc. Bras. de Computação

Barr, D. (1996). RFC 1912: Common DNS operational e configuration errors. http://www.ietf.org/rfc/rfc1912.txt.

Castro, S., Wessels, D., Fomenkov, M., e Claffy, K. (2008). A day at the root of the internet. *ACM SIGCOMM Computer Communication Review (CCR)*, 38(5):41–46.

Choi, H. e Lee, H. (2012). Identifying botnets by capturing group activities in dns traffic. *Computer Networks*, 56(1):20–33.

Crocker, D. (1997). RFC 2142: Mailbox names for common services, roles e functions. http://www.ietf.org/rfc/rfc2142.txt.

Dagon, D., Gu, G., Lee, C., e Lee, W. (2007). A taxonomy of botnet structures. In *Computer Security Applications Conference, 2007. ACSAC 2007. Twenty-Third Annual*, pages 325 –339.

DITL (2014). A day in the life of the internet (ditl). `https://www.dns-oarc.net/oarc/data/ditl` (acessado em 01/03/2014).

Ishibashi, K., Toyono, T., Toyama, K., Ishino, M., Ohshima, H., e Mizukoshi, I. (2005). Detecting mass-mailing worm infected hosts by mining dns traffic data. In *Proceedings of the 2005 ACM SIGCOMM workshop on Mining network data*, MineNet '05, pages 159–164, New York, NY, USA. ACM.

Jiang, N., Cao, J., Jin, Y., Li, L., e Zhang, Z.-L. (2010). Identifying suspicious activities through dns failure graph analysis. In *Network Protocols (ICNP), 2010 18th IEEE International Conference on*, pages 144–153.

Kumagai, M., Musashi, Y., Romana, D., Takemori, K., Kubota, S., e Sugitani, K. (2010). Ssh dictionary attack and dns reverse resolution traffic in campus network. In *Intelligent Networks and Intelligent Systems (ICINIS), 2010 3rd International Conference on*, pages 645–648.

Mockapetris, P. (1987). RFC 1034: Domain names - concepts and facilities. http://www.ietf.org/rfc/rfc1034.txt.

Ramachandran, A., Feamster, N., e Dagon, D. (2006). Revealing botnet membership using dnsbl counter-intelligence. In *Proceedings of the 2nd conference on Steps to Reducing Unwanted Traffic on the Internet - Volume 2*, SRUTI'06, pages 8–8, Berkeley, CA, USA. USENIX Association.

Shibata, N., Musashi, Y., Romana, D., Kubota, S., e Sugitani, K. (2012). Trends in host search attack in dns query request packet traffic. In *Intelligent Networks and Intelligent Systems (ICINIS), 2012 Fifth International Conference on*, pages 126–129.

Shin, S. e Gu, G. (2010). Conficker and beyond: A large-scale empirical study. In *Proceedings of the 26th Annual Computer Security Applications Conference*, ACSAC '10, pages 151–160, New York, NY, USA. ACM.

Wessels, D. e Fomenkov, M. (2003). Wow, that's a lot of packets. In *Passive and Active Network Measurement Workshop (PAM)*, pages 1–9, San Diego, CA. PAM.

Yuchi, X., Wang, X., Li, X., e Yan, B. (2009). Dns measurements at the .cn tld servers. In *Proceedings of the 6th international conference on Fuzzy systems and knowledge discovery - Volume 7*, FSKD'09, pages 540–545, Piscataway, NJ, USA. IEEE Press.

 ©2014 SBC — Soc. Bras. de Computação