PARTE I
INVESTIGAÇÃO DIGITAL

Esta parte do livro é especialmente dedicada aos operadores do Direito que possuam pouco ou nenhum conhecimento sobre as áreas de redes de computadores, segurança da informação, e para estudantes da área de Tecnologia da Informação que não saibam sobre a legislação relacionada à tecnologia. Serão abordados conceitos relacionados à investigação (rastreamento) através da Internet, fundamentos básicos de segurança da informação, *softwares* maliciosos (não apenas os sempre execrados vírus), alguns tipos de ataques a sistemas, noções do Direito aplicadas a crimes cibernéticos e, por fim, orientações seguras para a realização de busca e apreensão de equipamentos, além de um roteiro para a solicitação de perícia.

CAPÍTULO 1

Investigando Redes de Computadores

Um conceito simplificado para uma rede de computadores pode ser: dois ou mais dispositivos que utilizam um meio físico (cabo, ar, etc.) e protocolos (HTTP, SMTP, entre outros) para se comunicarem. Obviamente que este conceito abrange diversos dispositivos, não apenas computadores, que podem se conectar a uma rede, como por exemplo: *smartphones*, *blu-ray players*, televisores "inteligentes", entre outros.

Na atualidade, está ocorrendo uma migração do uso de gabinetes de computador (conhecidos por CPU ou *desktop*) para o uso de *notebooks*, *tablets*, *smartphones*, entre outros dispositivos portáteis. A principal finalidade para a aquisição destes dispositivos é a possibilidade de utilizar a Internet (rede mundial de computadores), através de pontos de acesso sem fio (*wireless access points*) ou através de modems 3G ou 4G.

Na prática, a maioria dos dispositivos conectados a uma rede utiliza a suíte de protocolos TCP/IP, que é o conjunto de protocolos utilizados na Internet, também adotado pela maioria das redes locais (LANs – *Local Area Networks*). Para que seja possível uma investigação em uma rede de computadores TCP/IP, é necessário que o investigador tenha conhecimento de como funciona: (1) o endereçamento das máquinas (endereço IP); (2) o serviço de resolução de nomes (DNS); (3) o serviço de tradução de endereços de rede (NAT); (4) um provedor de serviço Internet; (5) o rastreamento de um e-mail; (6) a busca pelo responsável por um endereço IP; (7) a busca de informações na Internet e (8) a coleta de informações com a devida validade jurídica. Nas seções que seguem, esses oito itens serão abordados.

1. Endereçamento IP

Um dos principais protocolos da suíte TCP/IP é o protocolo IP e a principal característica dele é o endereçamento, responsável pela identificação única de um dispositivo na rede. Se um dispositivo conectado à Internet estiver usando um determinado endereço IP, nenhum outro deverá utilizar o mesmo endereço IP no mesmo momento. Se por algum erro de configuração dois ou mais dispositivos utilizarem o mesmo endereço IP,

haverá conflito (geralmente detectado pelo sistema responsável) e dados podem ser entregues a mais de um destinatário, quando na verdade, deveriam ser entregues a apenas um.

Existem duas versões do protocolo IP: o IPv4, implantado em 1983; e o IPv6, implantado em 1999. O IPv6 surgiu da necessidade de mais endereços IP, pois com o crescente número de dispositivos ligados à Internet, havia uma estimativa que a quantidade de endereços IPv4 acabaria[1]. Outra maneira de postergar o término dos endereços IP disponíveis foi a utilização de endereços IP privados em redes locais e uma forma de traduzir estes em endereços válidos na Internet, através do serviço de tradução (NAT – ver Seção 3).

Um dispositivo pode possuir ambos endereços (IPv4 e IPv6), como é mostrado na Figura 1.1 através da ferramenta **ipconfig**. Como pode ser observado, o adaptador de rede sem fio possui o endereço IPv6 FE80::C9D3:AB99:2254:8B21%10 e o endereço IPv4 192.168.0.172.

Figura 1.1: Utilização da ferramenta ipconfig no Windows.

Na atualidade, o IPv4 ainda é o mais utilizado e quando o IPv6 estiver totalmente implementado em todos provedores do mundo, ainda assim o IPv4 continuará funcionando. Por esses motivos, o IPv4 será adotado neste livro e será chamado apenas de IP.

O endereço IP é formado por 32 bits, sendo representado por quatro números de 8 bits separados por ponto. Um número de 8 bits possui 256 valores diferentes ($2^8 = 256$), sendo o menor número 0 (00000000 em binário) e o maior 255 (11111111 em binário). É por esse motivo que nenhum dos quatro números é menor que 0 ou maior que 255 em um endereço IP.

1 O IPv4 permite 2^{32} endereços, enquanto o IPv6 permite 2^{128}. Mais detalhes podem ser obtidos em <http://ipv6.br/entenda/enderecamento/>.

Para organizar o endereçamento IP no mundo, faixas de endereçamento são distribuídas de forma hierárquica, sendo a IANA[2] (*Internet Assigned Numbers Authority*) a autoridade "raiz" responsável, tendo abaixo na hierarquia Registros de Internet Regionais (AfriNIC – África, APNIC – Ásia/Pacífico, ARIN – América do Norte, LACNIC – América Latina e algumas ilhas caribenhas, RIPE NCC – Europa, Ásia Central e Oriental). Pode haver registros nacionais, como por exemplo o CGI.br[3] (Comitê Gestor da Internet no Brasil). Desta forma, um provedor de serviço Internet (ISP – *Internet Service Provider*) deve contratar faixa(s) de endereços IP através dos registros regionais (ou nacionais, quando houver) e, quando um cliente do provedor se conectar à Internet, deverá receber um dos endereços IP pertencentes à faixa mencionada (um endereço IP disponível, que não estiver em uso por outro cliente).

2. Serviço de Resolução de Nome de Domínio (DNS)

Como já foi descrito na Seção 1, os dispositivos são identificados na Internet através do endereço IP, porém seres humanos têm dificuldade em memorizar diversos endereços numéricos para acessar páginas, e-mails, servidores de arquivos e outros serviços. Uma solução encontrada foi a criação de um serviço de tradução de nomes para os endereços IP equivalentes, o *Domain Name Service* (DNS). A delegação de domínios de mais alto nível (*top-level domain* - TLD), tais como .com, .edu, .br, .uk, .uy, entre outros, é de responsabilidade da IANA[4]. Para o Brasil (TLD .br), a IANA define como responsável o CGI.br, ou seja, para registrar um domínio com "final" .br deve-se contratar somente através do CGI.br. Por exemplo, se alguém quiser registrar o domínio evandrodellavecchia.com.br, deve verificar se há a disponibilidade deste domínio, através do sítio <http://registro.br>, e se houver, realizar a solicitação, efetuar o pagamento e informar as configurações solicitadas pelo CGI.br sobre o provedor onde o sítio será hospedado.

Na medida em que novos domínios são cadastrados, estes são propagados pela Internet e em poucas horas todos servidores DNS do mundo são capazes de traduzir o domínio para o endereço IP equivalente onde está hospedado o sítio (ou outro serviço qualquer). Para

2 Sítio da IANA disponível em <http://www.iana.org/>.

3 Sítio do CGI.br disponível em <http://www.cgi.br>.

4 Base de dados de domínios de mais alto nível (*top-level domain*) disponível em <http://www.iana.org/domains/root/db>.

não haver consultas constantes a servidores DNS de mais alto nível (mais próximos da raiz, ou a própria raiz), os servidores DNS geralmente possuem uma memória *cache*[5], respondendo imediatamente ao solicitante (quando tiver a resposta), poupando tempo de espera e tráfego gerado na Internet. É possível também, em sistemas operacionais como Windows e Linux, configurar em traduções fixas, de domínio para endereço IP, como por exemplo: evandrodellavecchia.com.br para 184.172.11.186.

3. Tradução de Endereçamento de Rede (NAT)

Quando o protocolo IPv4 foi criado, não se imaginava que haveria um crescimento capaz de tornar o endereçamento IP escasso. Uma estratégia adotada foi a utilização de endereçamento privado[6] (não roteável). As faixas reservadas para o endereçamento privado são mostradas no Quadro 1.1.

10.0.0.0 a 10.255.255.255
172.16.0.0 a 172.31.255.255
192.168.0.0 a 192.168.255.255

Quadro 1.1: Faixas de endereçamento IP privado.

Dispositivos que utilizam endereço IP dentro das faixas mostradas no Quadro 1.1 não podem se comunicar diretamente na Internet, necessitando que uma tradução de endereçamento seja realizada. O serviço responsável por esta tradução é o *Network Address Translation* (NAT). Geralmente este serviço fica habilitado em roteadores ou modems/roteadores. Para um melhor entendimento, um cenário será mostrado a seguir.

Supondo uma rede local com quatro equipamentos, sendo dois ligados com cabo e dois utilizando conexão sem fio a um equipamento roteador *wireless* D-Link DI-524[7], e o roteador ligado a um *cable* modem (Figura 1.2), tem-se a configuração mostrada no Quadro 1.2.

5 Responsável por armazenar consultas recentes, respondendo ao solicitante diretamente, sem ter que consultar servidores de mais alto nível.

6 RFC 1918 - *Address Allocation for Private Internets*. Disponível em <http://www.rfc-editor.org/rfc/rfc1918.txt>.

7 Detalhes do produto disponíveis em <http://www.dlink.com.br/produtos-detalhes/items/di-524.html>.

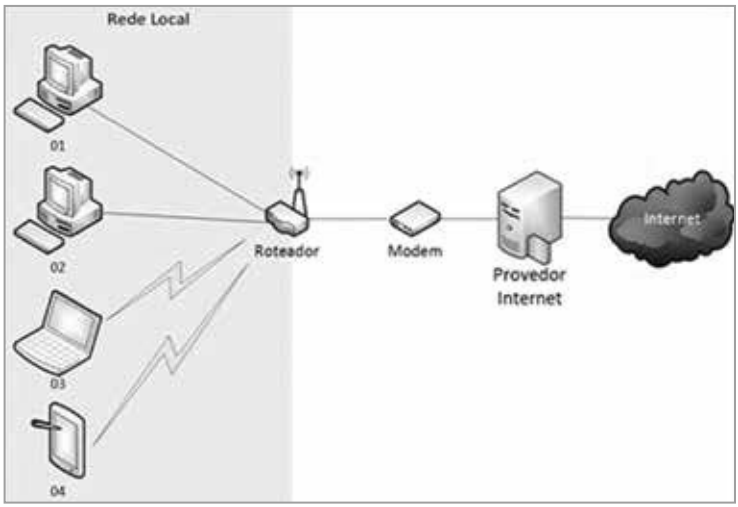

Figura 1.2: Cenário de uma rede local ligada à Internet.

Roteador:
Endereço IP da interface da rede local: 192.168.0.1 (padrão do modelo de roteador apresentado)
Endereço IP da interface ligada ao modem: 189.6.138.222 (exemplo de endereço IP recebido pelo provedor Internet)
Dispositivos na rede local:
01: 192.168.0.171
02: 192.168.0.172
03: 192.168.0.173
04: 192.168.0.174

Quadro 1.2: Endereçamento IP dos equipamentos mostrados na Figura 1.2.

Por exemplo, se um usuário utilizar o Dispositivo 02 para acessar uma página Web, a solicitação será encaminhada ao *gateway* (interface da rede local do roteador), que traduzirá o endereço 192.168.0.172 para 189.6.138.222 e encaminhará a solicitação ao servidor da página solicitada. O retorno será do servidor ao roteador (endereço IP externo), que traduzirá para o endereço 192.168.0.172 e enviará ao Dispositivo 02.

Na prática, a grande maioria das empresas utiliza o esquema mostrado no cenário apresentado. Isto quer dizer que se alguém cometer algum ilícito através da Internet a partir da rede local de uma empresa e esta utilizar NAT, a investigação terá condições de realizar o rastreamento até a empresa, mas (geralmente) não terá condições de saber qual das máquinas foi utilizada. A empresa pode colaborar com a investigação, caso ela tenha registros de acessos à Internet e mapeamento dos seus dispositivos com seus devidos

endereço IP. Caso a empresa não possua tais registros, pode responder judicialmente. Este é um motivo pelo qual estabelecimentos como hotéis, restaurantes, entre outros, começaram a exigir a autenticação de seus clientes (através de cadastro prévio), mesmo quando a Internet não é cobrada.

4. Provedor de Acesso à Internet (ISP)

O provedor de acesso à Internet (*Internet Service Provider* - ISP) oferece como principal serviço o acesso à Internet, podendo oferecer também outros serviços, tais como hospedagem de sítios, contas de e-mail, repositório de dados, entre outros. Para poder utilizar a Internet, o dispositivo que interliga a rede local ao ISP deve se autenticar. Atualmente, é comum que o próprio ISP realize a autenticação, mas é possível utilizar um canal de comunicação de uma operadora e a autenticação de um ISP. Por exemplo, pode-se utilizar linha telefônica de uma determinada operadora como meio de comunicação e a autenticação pode ser realizada por outra empresa.

O registro destas autenticações é de extrema importância para a investigação, pois através dele é possível identificar que um cliente iniciou uma conexão em determinada data/horário, qual endereço IP recebeu, e a data/horário de desconexão. Em uma nova conexão, o mesmo usuário pode receber outro endereço IP, por isso é de suma importância que a investigação saiba além do endereço IP, o momento exato que uma atividade ilícita tenha ocorrido.

5. Rastreamento de E-mail

A correspondência eletrônica (e-mail) é um dos serviços mais utilizados na Internet e basicamente possui um emissor, um ou mais destinatários, o assunto e o conteúdo da mensagem. Estas são as informações mostradas ao usuário, tanto em softwares cliente de e-mail (Microsoft Outlook, Incredimail, entre outros) como em serviços de e-mail disponibilizados diretamente em um sítio (*Webmail*). Porém, além das informações mencionadas, outras compõem o e-mail, como o endereço IP origem, o caminho percorrido entre o emissor e o destinatário, entre outras. Estas informações são adicionadas conforme ocorre o envio em cada servidor de e-mail pelo qual passa. Elas podem ser localizadas no cabeçalho do e-mail e, para serem exibidas, deve-se clicar em algum botão ou menu, dependendo do software cliente ou *Webmail* utilizado. A RFC 2821[8] descreve os campos do cabeçalho do SMTP (*Simple Mail Transfer Protocol*), protocolo utilizado para envio de e-mails.

8 *Request for Comments* (RFC) 2821 – Padrão SMTP. Disponível em <http://www.ietf.org/rfc/rfc2821.txt>.

Para um melhor entendimento de como coletar e compreender o cabeçalho de um e-mail, foram realizados envios de e-mail entre quatro servidores de e-mail, sendo dois serviços gratuitos: Gmail e Hotmail; e dois privados: PUCRS e Procergs. Para a leitura dos e-mails, foram utilizados dois softwares: Microsoft Outlook e Incredimail; e acesso direto a dois Webmails: Gmail e Hotmail. Serão mostradas telas dos softwares e/ou navegadores e os cabeçalhos dos e-mails recebidos (os nomes das contas foram alteradas para não serem identificadas nas figuras, utilizando os nomes "remetente" e "destinatario" ao invés dos nomes reais).

5.1. Primeiro experimento:

No primeiro experimento foi enviado um e-mail de uma conta remetente@pucrs.br para destinatario@gmail.com. Para a leitura do e-mail, foi utilizado o software Microsoft Outlook. Na Figura 1.3 é possível observar que para visualizar o cabeçalho há o menu "Opções de Mensagem...", que entre outras informações mostra o cabeçalho (Figura 1.4). O cabeçalho completo do e-mail selecionado na Figura 1.4 é mostrado no Quadro 1.3, sendo que os nomes dos campos foram formatados em negrito.

Figura 1.3: Menu para mostrar o cabeçalho do e-mail no Microsoft Outlook.

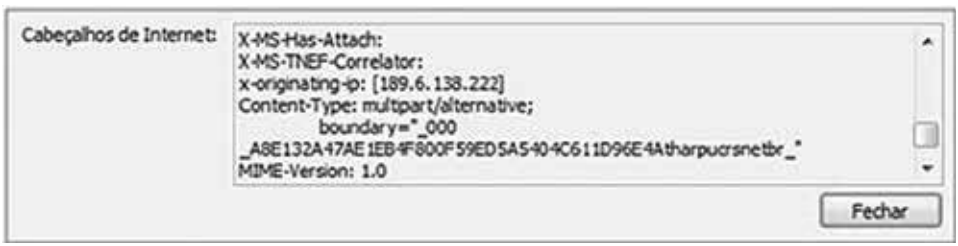

Figura 1.4: Trecho do cabeçalho do e-mail.

Delivered-To: destinatario@gmail.com
Received: by 10.182.241.232 with SMTP id wl8csp66578obc;
Sun, 18 Aug 2013 17:13:51 -0700 (PDT)
X-Received: by 10.236.30.41 with SMTP id j29mr10988740yha.0.1376871230943;
Sun, 18 Aug 2013 17:13:50 -0700 (PDT)
Return-Path: <remetente@pucrs.br>
Received: from nash.pucrs.br (nash.pucrs.br. [201.54.140.129])
by mx.google.com with ESMTP id j62si202929yhl.140.1969.12.31.16.00.00;
Sun, 18 Aug 2013 17:13:50 -0700 (PDT)
Received-SPF: pass (google.com: domain of remetente@pucrs.br designates 201.54.140.129 as permitted sender) client-ip=201.54.140.129;
Authentication-Results: mx.google.com;
spf=pass (google.com: domain of remetente@pucrs.br designates 201.54.140.129 as permitted sender) smtp.mail=remetente@pucrs.br
Received: from gobi.pucrsnet.br (Not Verified[10.40.19.18]) by nash.pucrs.br with MailMarshal (v6,9,6,3437) id <B5211633d0000>; Sun, 18 Aug 2013 21:13:49 -0300
Received: from THAR.pucrsnet.br ([169.254.2.151]) by gobi.pucrsnet.br
([10.40.19.18]) with mapi id 14.02.0318.004; Sun, 18 Aug 2013 21:13:49 -0300
From: Nome do Remetente <remetente@pucrs.br>
To: “destinatario@gmail.com” <destinatario@gmail.com>
Subject: URGENTE!
Thread-Topic: URGENTE!
Thread-Index: Ac6ccPoSXUEe/IeMRV+ngIBU6TKb/g==
Date: Mon, 19 Aug 2013 00:13:48 +0000
Message-ID: <A8E132A47AE1EB4F800F59ED5A5404C611D96E4A@thar.pucrsnet.br>
Accept-Language: pt-BR, en-US
Content-Language: pt-BR
X-MS-Has-Attach:
X-MS-TNEF-Correlator:
x-originating-ip: [189.6.138.222]
Content-Type: multipart/alternative;
boundary=”_000_A8E132A47AE1EB4F800F59ED5A5404C611D96E4Atharpucrsnetbr_”
MIME-Version: 1.0

Quadro 1.3: Cabeçalho completo do e-mail do primeiro experimento.

Seguindo a especificação RFC 2821, pode-se concluir que o e-mail mostrado foi enviado:

- de (From:) Nome do remetente <remetente@pucrs.br>;

- para (To:) destinatario@gmail.com <destinatario@gmail.com>;
- no dia (Date:) 19/08/2013 às 0h13min48s (horário GMT – Greenwich Mean Time, pois mostra +0000[9]), ou seja, foi enviado no dia 18/08/2013 às 21h13min48s no horário de Brasília;
- com o assunto (Subject:) URGENTE!;
- com endereço de retorno (Return-Path:) remetente@pucrs.br;
- com informações de servidores que receberam e encaminharam o e-mail (Received:), tais como: nomes dos servidores (THAR.pucrsnet.br, gobi.pucrsnet.br, nash.pucrs.br, mx.google.com), seus endereços IP, entre outras.

Alguns campos encontrados no cabeçalho não são especificados na RFC 2821, mas são utilizados por alguns softwares e/ou Webmails. Um exemplo é o endereço IP do dispositivo que redigiu o e-mail (x-originating-ip:[10]), campo criado pela Microsoft para inibir que usuários utilizem o serviço Hotmail para enviar *spam*[11]. Analisando o campo x-originating-ip:, pode-se observar o endereço IP 189.6.138.222.

Resumindo, um dispositivo com endereço IP 189.6.138.222 redigiu um e-mail (remetente@pucrs.br), que foi enviado através do servidor de e-mail com endereço IP 201.54.140.129, para o e-mail destinatario@gmail.com. Na Seção 6 será mostrado como descobrir o responsável pelos endereços IP citados tanto neste primeiro experimento como nos que seguem.

5.2. Segundo experimento:

No segundo experimento foi enviado um e-mail de uma conta remetente@gmail.com para destinatario@pucrs.br. Para a leitura do e-mail, foi utilizado o Webmail Outlook WebApp. Através da sequência de menus/botões: Ações, Exibir mensagem original, Detalhes da mensagem, foi possível visualizar o cabeçalho. Trechos do cabeçalho são mostrados no Quadro 1.4 (trechos suprimidos no referido Quadro e em outros que aparecem nesta Seção foram substituídos por [...], para uma melhor visualização).

9 Se estivesse representado com horário GMT com relação a Brasília, seria mostrado -0300, ou seja, três horas a menos que GMT. E se fosse durante o horário de verão, seria mostrado -0200. Mais informações podem ser obtidas em <http://wwp.greenwichmeantime.com/>.

10 Mais informações disponíveis em <http://www.microsoft.com/en-us/news/features/1999/09-22spam.aspx>.

11 Mensagem com conteúdo não solicitado, geralmente indesejado, como por exemplo propagandas.