EspiaMule e Wyoming ToolKit: Ferramentas de Repressão à Exploração Sexual Infanto-Juvenil em Redes Peer-to-Peer

Jorge Ricardo Souza de Oliveira e Edmar Edilton da Silva, *Peritos Criminais Federais, Departamento de Polícia Federal*

***Resumo* – Este artigo apresenta as ferramentas EspiaMule e Wyoming Toolkit que são utilizadas para auxiliar na identificação dos responsáveis pela disponibilização de material de exploração sexual infanto-juvenil em redes de compartilhamento de arquivos do tipo peer-to-peer. Ademais, é apresentado um estudo de caso de utilização dessas ferramentas, levando em consideração a situação brasileira em relação à divulgação desse tipo de material.**

***Palavras-chave* – P2P, peer-to-peer, eDonkey2000, Kad, Gnutella, EspiaMule, Wyoming ToolKit, WTK, pedofilia e perícia forense.**

I. INTRODUÇÃO

A exploração sexual infanto-juvenil é um dos problemas que, atualmente, mais preocupa a sociedade brasileira e mundial. Apesar desse tipo de exploração consistir em um problema antigo, a Internet tem fornecido novos meios para a publicação e distribuição de material a este respeito, aumentando, em muito, o seu alcance. Entre esses novos meios, podem ser citados o correio eletrônico, as redes sociais, as ferramentas de comunicação instantânea e as redes de compartilhamento de arquivos do tipo peer-to-peer.

Em especial, o desenvolvimento das redes de compartilhamento de arquivos do tipo peer-to-peer possibilitou aos usuários da Internet a troca de arquivos sem a necessidade de servidores centralizados [3, 4, 6]. Tal característica, além de, potencialmente, aumentar as taxas de transmissão de dados, já que é eliminada a necessidade de um servidor central, também cria dificuldades na identificação dos responsáveis pela disponibilização ou distribuição indevida de certos tipos de arquivos. Entre esses arquivos encontram-se material protegido por direitos autorais distribuídos sem autorização e material relacionado a outros ilícitos penais como discriminação, calúnia, difamação, injúria, estelionato, incitação e apologia ao crime e exploração sexual infanto-juvenil.

É importante observar que existem diversas redes de compartilhamento de arquivos do tipo peer-to-peer. Segundo dados de pesquisa desenvolvida pelo IBOPE/Netrating [9], em setembro de 2008, 42,9% dos internautas residenciais brasileiros usaram aplicativos de compartilhamento de arquivos do tipo peer-to-peer. Segundo essa mesma pesquisa, os aplicativos mais populares são o eMule, o Limewire e o Ares Galaxy, sendo utilizados, respectivamente, por 14,2%, 8,5% e 8,4% dos internautas residenciais brasileiros. Enquanto o aplicativo eMule utiliza as redes eDonkey2000 e Kad, os aplicativos Limewire e Ares Galaxy utilizam, respectivamente, as redes Gnutella e Ares.

A identificação dos responsáveis pela disponibilização de material de exploração sexual infanto-juvenil é uma atividade complexa e, geralmente, envolve uma série de questões. As ferramentas EspiaMule e Wyoming ToolKit (WTK) auxiliam nessa identificação fornecendo, entre outras informações, o endereço IP (*Internet Protocol*) e o identificador do aplicativo (*hash* do aplicativo) utilizados pelos computadores que estão disponibilizando esse material nas redes eDonkey2000, Kad e Gnutella.

Este artigo tem como objetivo apresentar as principais características das ferramentas EspiaMule e WTK, além de propor um estudo de caso para cada uma das ferramentas na tentativa de identificar qual delas melhor se aplica, levando em consideração a casuística no Brasil em relação à divulgação de material de exploração sexual infanto-juvenil nas redes de compartilhamento de arquivos do tipo peer-to-peer [5].

II. ESPIAMULE

A ferramenta EspiaMule foi idealizada e desenvolvida pelos Peritos Criminais Federais Guilherme Martini Dalpian e Carlos Augustus Armelin Benites do Departamento de Polícia Federal do Brasil. O EspiaMule foi criado a partir do código fonte original do aplicativo eMule com a finalidade de realizar

Jorge Ricardo Souza de Oliveira é Perito Criminal Federal do Departamento de Polícia Federal, lotado no Setor Técnico Científico da Superintendência Regional de Polícia Federal no Estado do Paraná. E-mail: jorge.jrso@dpf.gov.br.

Edmar Edilton da Silva é Perito Criminal Federal do Departamento de Polícia Federal, lotado no Setor Técnico Científico da Superintendência Regional de Polícia Federal no Estado do Paraná. E-mail: edmar.ees@dpf.gov.br.

o monitoramento das redes eDonkey e Kad utilizadas por este aplicativo [1, 7].

Na implementação do EspiaMule, o código fonte do aplicativo eMule foi modificado com os seguintes objetivos:

1) Localizar nas redes eDonkey e Kad os endereços IP que estejam compartilhando materiais ilícitos, como, por exemplo, pornografia infanto-juvenil;
2) Não permitir o envio (*upload*) de arquivos para outros usuários da rede.

A Figura 1 mostra a interface inicial do aplicativo EspiaMule.

Fig. 1. Interface inicial do aplicativo EspiaMule.

O Espiamule pode ser utilizado para identificar computadores na Internet que estejam compartilhando qualquer tipo de arquivo relacionado a ilícitos penais. Utilizando o EspiaMule pode ser feito o *download* de arquivos a partir de *links* ED2K ou buscas por palavras-chave contidas nos nomes dos arquivos procurados.

Ao se realizar o *download* de um arquivo a partir de seu *link* ED2K, há garantia de que os computadores localizados na rede, realmente, compartilham o arquivo ou parte dele. Esta garantia não existe no caso de buscas por palavras-chave, já que os arquivos localizados podem conter os termos pesquisados, mas os seus conteúdos não estarem relacionados às buscas realizadas.

Um *link* ED2K identifica um arquivo na Internet de forma única através de seu código *hash* (gerado pelo algoritmo MD4). Um *link* ED2K para um arquivo apresenta o seguinte formato:

ed2k://|file|nome do arquivo|tamanho|código hash|/

O campo "*nome do arquivo*" representa o nome do arquivo com sua extensão (exemplo "Exemplo.jpg"), o campo "*tamanho*" representa o tamanho do arquivo em bytes (exemplo 97525) e o código *hash* é o identificador do arquivo (exemplo 34008ED7063170DCA3707EA59B48CE59).

Para possibilitar um melhor controle sobre os arquivos que estão sendo baixados (*download*) da Internet, o aplicativo EspiaMule divide-os em partes de aproximadamente 9 MB. Assim, o código *hash* de um arquivo que foi totalmente baixado é a combinação do código *hash* de cada uma das partes desse arquivo.

Para localizar na Internet computadores que estejam compartilhando arquivos, por exemplo, relacionados à pornografia infanto-juvenil, basta gerar uma lista contendo os *links* ED2K destes arquivos e utilizar a ferramenta EspiaMule para realizar o *download* desses arquivos. O EspiaMule irá criar automaticamente um arquivo de *log* (formato csv) contendo o endereço IP, *hash* do aplicativo, nome do país, nome do arquivo, *hash* do arquivo, data e hora dos computadores compartilhando os arquivos na rede. O fuso horário utilizado na hora é o horário brasileiro (sem horário de verão, GMT -0300).

Ao ser instalado em um computador, o EspiaMule gera automaticamente um *hash* que identifica o aplicativo daquela instalação de forma única na rede. O *hash* do aplicativo é persistente entre várias conexões sendo substituído somente no caso de uma nova instalação do aplicativo.

O arquivo de *log* pode ser processado utilizando um programa em Java, denominado eMuleWhoisParser, que separa automaticamente os usuários por provedor de acesso através de pesquisas a serviços do tipo WHOIS para os endereços IP localizados no Brasil e agrupa os usuários com diversos arquivos compartilhados a partir do *hash* do aplicativo ou pelo endereço IP utilizado.

Ressalta-se que o nome de um arquivo compartilhado nas redes eDonkey e Kad é escolhido de forma arbitrária e pode não corresponder a seu conteúdo. Dessa forma, como muitos usuários dessas redes fazem *downloads* de arquivos a partir de buscas por termos contidos em seus nomes desses arquivos, é possível que sejam baixados arquivos com conteúdo diferente do desejado. Assim, é importante que a utilização do aplicativo Espiamule seja focada em usuários que compartilhem vários arquivos com conteúdo ilícito (por exemplo, cinco ou mais arquivos). Isto tenderia a diminuir a quantidade de usuários identificados pelo EspiaMule que, porventura, teriam feito o *download* de arquivos com conteúdo diferente do desejado.

III. WYOMING TOOLKIT (WTK)

A ferramenta WTK foi desenvolvida no contexto da Operação Fairplay que consiste em um esforço cooperativo internacional desenvolvido por diversas instituições policiais espalhadas no mundo e tem o intuito de combater crimes relacionados à divulgação e distribuição de material de exploração sexual infanto-juvenil na Internet. Em especial, esta ferramenta permite localizar e documentar informações sobre arquivos disponibilizados na rede de compartilhamento de arquivos do tipo peer-to-peer denominada Gnutella [2].

O pacote WTK consiste em uma série de aplicativos e uma base de dados mantida pela Divisão de Investigações Criminais

do Estado de Wyoming nos Estados Unidos da América. Atualmente, essa base de dados encontra-se em processo de replicação, sendo que cópias da mesma devem ser mantidas no Canadá e na Austrália. Os usuários licenciados do pacote WTK podem registrar informações de suas investigações e acessar dados de investigações empreendidas por outras instituições policiais. Entre as informações mantidas na base de dados, encontram-se uma lista de códigos *hash* de arquivos relacionados à exploração sexual infanto-juvenil e dados históricos sobre endereços IP e códigos *hash* de aplicativos que foram identificados, em algum momento, como detentores desses arquivos. Um *hash* de aplicativo consiste em um valor que identifica unicamente um cliente da rede Gnutella. É importante ressaltar que, como não há um controle a respeito dos dados repassados das investigações, tais dados devem ser usados apenas como indicativos e necessitam de confirmação de cada usuário que utiliza o pacote WTK.

Entre as ferramentas que compõe o pacote WTK estão o WTK-cliente, desenvolvido por Flint Waters, Agente Especial da Divisão de Investigações Criminais do Estado de Wyoming, bem como, o GnuWatch e o Peer Spectre, desenvolvidos por William Wiltse, integrante do Departamento de Polícia de Salem no Oregon.

A. *WTK-cliente*

A ferramenta WTK-cliente deve ser usada em conjunto com um aplicativo cliente da rede Gnutella e permite o registro e o acesso a informações da base de dados do pacote WTK. Embora a sua utilização não seja obrigatória, o WTK-cliente foi otimizado para uso em conjunto com o aplicativo Phex, um cliente da rede Gnutella de código aberto. Por meio do cliente Gnutella escolhido, deve ser feita uma busca por termos relacionados à exploração sexual infanto-juvenil, sendo retornada uma lista de nomes de arquivos que contém esses termos e os códigos *hash* desses arquivos. Então, o WTK-cliente deve ser utilizado para identificar quais códigos *hash* obtidos são cadastrados como pertencentes a arquivos de exploração sexual infantil. Em seguida, o cliente Gnutella deve ser utilizado para identificar os computadores dessa rede que estão fornecendo os arquivos com os códigos *hash* identificados no passo anterior. De cada computador identificado é recuperado o seu endereço IP e o *hash* do aplicativo cliente da rede Gnutella que está sendo executado. Por fim, o WTK-cliente deve ser utilizado para cadastrar os endereços IP recuperados e verificar a possível localização física aproximada desses endereços. As informações de localização retornadas incluem a latitude e longitude e, muitas vezes, a cidade, o estado e o país. Tais informações são especialmente importantes para a realização de investigações restritas a apenas uma região [2]. A Figura 2 mostra a interface inicial do aplicativo WTK-cliente.

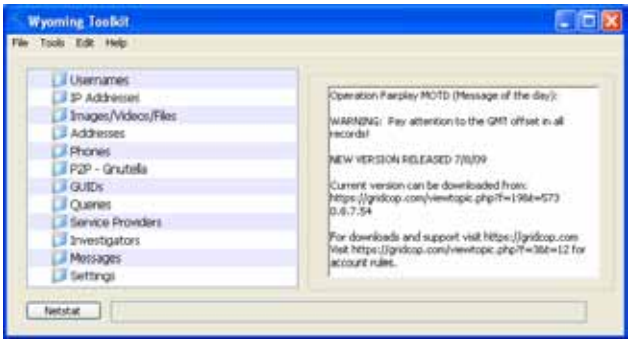

Fig. 2. Interface inicial do aplicativo WTK-cliente.

B. *Peer Spectre*

A ferramenta Peer Spectre realiza, continuamente, buscas na rede Gnutela e identifica computadores dessa rede que contêm arquivos de exploração sexual infanto-juvenil. Dessa forma, o trabalho desenvolvido em conjunto pelo WTK-cliente e pelo aplicativo cliente da rede Gnutella, descrito na seção anterior, é realizado sem a necessidade de interação constante com o usuário. Os endereços IP dos computadores identificados pela ferramenta Peer Spectre são retornados em um arquivo de log [2]. A Figura 3 mostra a interface inicial do aplicativo Peer Spectre.

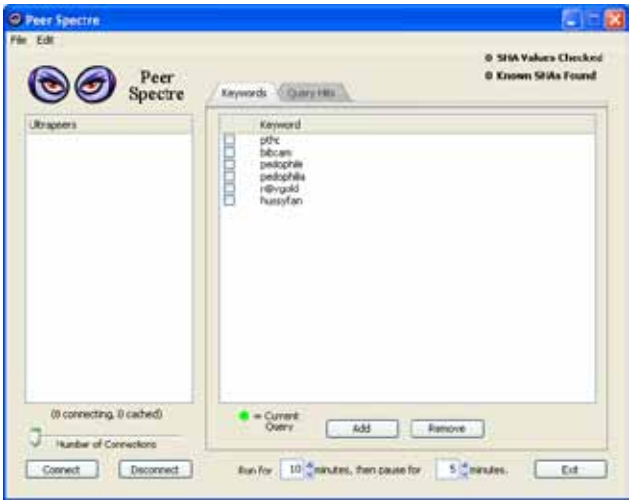

Fig. 3. Interface inicial do aplicativo Peer Spectre.

C. *GnuWatch*

A ferramenta GnuWatch permite o monitoramente contínuo de um conjunto de endereços IP previamente identificados por outras ferramentas. Por meio do GnuWatch é possível verificar quando determinados computadores que possuem os endereços IP identificados estão conectados na rede Gnutella, bem como baixar os arquivos disponibilizados por esses computadores [2]. A Figura 4 mostra a interface inicial do aplicativo GnuWatch.

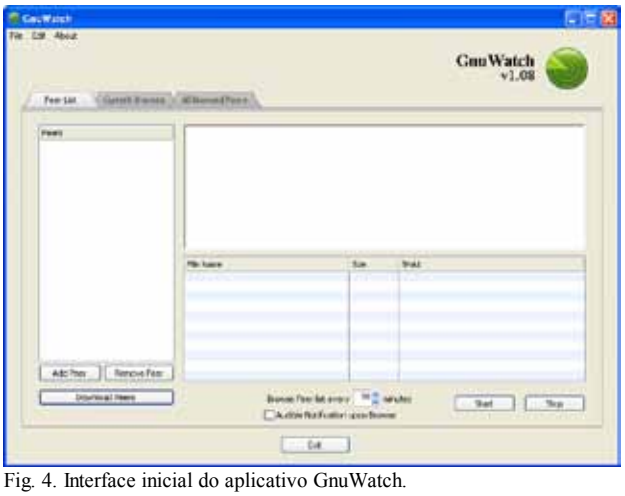

Fig. 4. Interface inicial do aplicativo GnuWatch.

IV. Materialidade e Autoria

As ferramentas EspiaMule e WTK fornecem o endereço IP e o identificador do aplicativo que estão compartilhando arquivos pesquisados. Essas duas informações são essenciais no estabelecimento da materialidade do ilícito penal e na tentativa de identificar sua autoria.

A. Materialidade

A materialidade pode ser comprovada por meio da localização de arquivos que estão sendo compartilhados com outros usuários da rede. Todavia, ressalta-se que é necessário ter certeza que os arquivos compartilhados possuem, realmente, conteúdo que caracterize um ilícito penal, como, por exemplo, exploração sexual infanto-juvenil.

B. Autoria

A autoria do ilícito penal não é fácil de ser identificada, pois encontrar o endereço IP e o *hash* do aplicativo utilizado no compartilhamento dos arquivos não significa identificar as pessoas responsáveis pelo compartilhamento, haja visto que o computador utilizado pode estar localizado, por exemplo, em uma *lan house*, empresa ou residência onde várias pessoas compartilham o mesmo computador.

Outros fatores que dificultam a identificação dos responsáveis pelos compartilhamentos são a utilização de:

1) NAT (*Network Address Translation*).
2) Endereço IP dinâmico.
3) Proxy.

Nesses casos, normalmente, é necessário um pedido de quebra de sigilo telemático para que, a partir do endereço IP, bem como data e hora (fuso horário), seja solicitado ao provedor de acesso à Internet o fornecimento do endereço e do responsável pelo terminal telefônico ou de dados que utilizou o IP na data e hora informadas. Outra dificuldade encontrada na identificação da autoria do ilícito penal é que o *hash* do aplicativo é passível de ser clonado.

Apesar das dificuldades mencionadas, uma vez que um usuário da rede usando um determinado computador aparece nas relações de endereços IP geradas pelas ferramentas EspiaMule e WTK, significa que esse usuário compartilha ou compartilhou com outros usuários da rede pelo menos um dos arquivos pesquisados (ou parte do arquivo). Assim, a partir da localização física do computador que compartilhou os arquivos, o *hash* do aplicativo pode ser utilizado na tentativa de estabelecer a autoria do ilícito penal.

A confirmação de que o *hash* do aplicativo no computador localizado correponde a um *hash* monitorado pode ainda não ser suficiente para comprovar a autoria. Uma perícia nesse computador e uma investigação apurada podem ser necessárias para se chegar às pessoas responsáveis pelo compartilhamento dos arquivos.

V. Estudo de Caso

Nesta seção é proposto um estudo de caso para as ferramentas EspiaMule e WTK com a finalidade de identificar qual delas melhor se aplica à casuística no Brasil em relação à divulgação de material de exploração sexual infanto-juvenil nas redes de compartilhamento de arquivos peer-to-peer.

Sobre os resultados expostos a seguir, vale ressaltar que a localização dos endereços IP identificados pelas ferramentas foi realizada de forma automatizada.

A. Estudo de Caso 1 (EspiaMule)

Conforme explicado na Seção II, o EspiaMule utiliza *links* ED2K para fazer o *download* de arquivos e registrar as informações que possibilitam a identificação do computador responsável pelo compartilhamento. Como o foco do artigo é direcionado para a repressão à exploração sexual infanto-juvenil, faz-se necessária a utilização de um conjunto de códigos *hash* gerados a partir de arquivos contendo material relacionado à exploração sexual infanto-juvenil. Neste estudo de caso, foram utilizados 131 *links* ED2K referentes a arquivos comprovadamente relacionados à exploração sexual infanto-juvenil.

No dia 14 de julho de 2009, às 17h30min, foi iniciado o processo de busca nas redes eDonkey e Kad pelos computadores que estavam, naquele momento, compartilhando pelo menos parte de um dos arquivos contidos na lista de *links* ED2K pesquisados. O processo de buscas foi finalizado às 10h02min do dia 16 de julho de 2009. Os resultados obtidos foram registrados em um arquivo de *log*.

Foi utilizado o programa eMuleWhoisParser, com o arquivo de *log* gerado pelo EspiaMule, para retirar as ocorrências de endereços IP duplicados e identificar os endereços IP localizados no Brasil. O programa eMuleWhoisParser possibilita que os dados gerados sejam agrupados por provedor de acesso, facilitando o processo de investigação iniciado após a obtenção dos endereços IP. No Brasil, foram encontrados 118 diferentes provedores de acesso à Internet ou empresas que contratam acesso à Internet com endereço IP dedicado.

Analisando o arquivo de *log*, também foi verificado que 42.814 endereços IP, em 107 países, compartilham pelo menos parte de um dos arquivos pesquisados. No Brasil, foram encontrados 1.120 endereços IP e 997 códigos *hash* de aplicativos compartilhando os arquivos pesquisados. Nas Figuras 5 e 6, são ilustrados os países onde estavam localizadas as maiores quantidades de endereços IP e códigos *hash* de aplicativos.

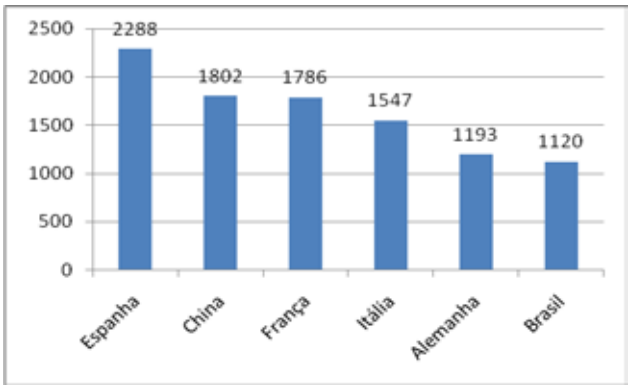

Fig. 5. Localização de endereços IP identificados pelo EspiaMule.

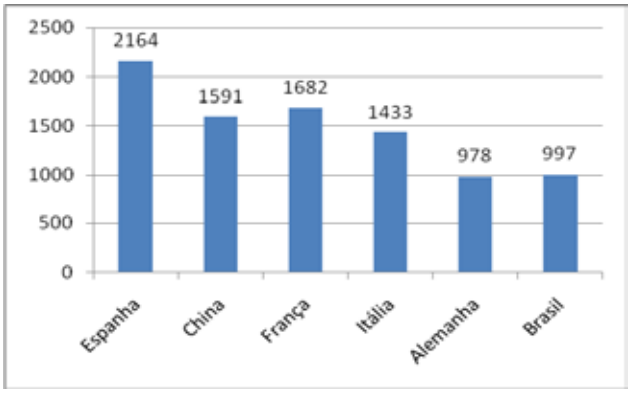

Fig. 6. Localização de códigos *hash* de aplicativo identificados pelo EspiaMule.

No que diz respeito aos aplicativos localizados no Brasil, a maior parte estava compartilhando apenas um dos arquivos pesquisados. Na Figura 7, é ilustrado o número de aplicativos localizados no Brasil compartilhando arquivos.

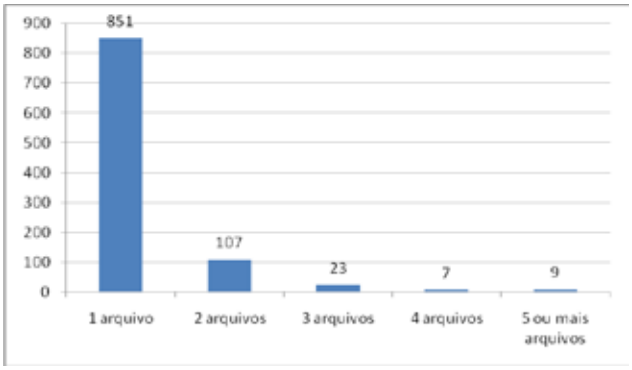

Fig. 7. Número de aplicativos localizados no Brasil compartilhando arquivos.

B. Estudo de Caso 2 (WTK)

Para a realização do estudo de caso do pacote WTK, a ferramenta Peer Spectre foi executada das 9h00min do dia 16 de julho de 2009 até as 9h00min do dia seguinte. Essa ferramenta foi utilizada por não exigir interação constante com o usuário. Durante o período citado, a ferramenta Peer Spectre realizou buscas por arquivos cujos nomes contêm termos relacionados à exploração sexual infantil e comparou os códigos *hash* dos arquivos encontrados com uma base de dados local de códigos *hash* de arquivos conhecidos. Os resultados da execução foram armazenados em um arquivo de *log*.

No período de execução, a ferramenta Peer Spectre localizou 479 arquivos relacionados à exploração sexual infanto-juvenil com códigos *hash* conhecidos. Tais arquivos estavam sendo fornecidos por 310 endereços IP diferentes. Conforme ilustrado na Figura 8, a maior parte dos endereços IP localizados encontrava-se nos Estados Unidos da América. Ressalta-se que nenhum desses endereços IP encontrava-se no Brasil.

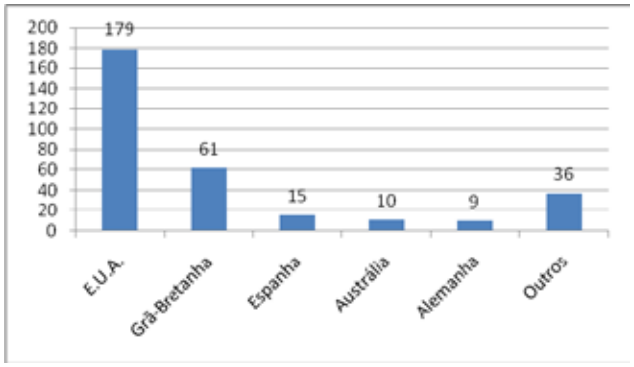

Fig. 8. Localização dos endereços IP identificados pelo Peer Spectre.

Observou-se ainda que, dos 310 endereços IP identificados, 137 estavam compartilhando apenas 1 arquivo e 79 estavam compartilhando 5 ou mais arquivos. Ademais, dos 79 endereços IP compartilhando 5 ou mais arquivos, 55 estavam localizados nos Estados Unidos da América. Mais informações são ilustradas nas Figuras 9 e 10.

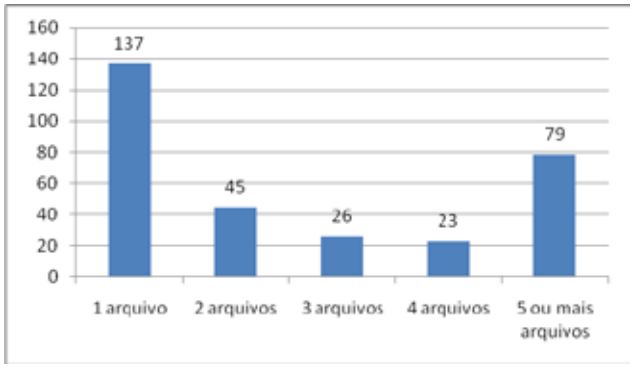

Fig. 9. Número de endereços IP compartilhando arquivos.

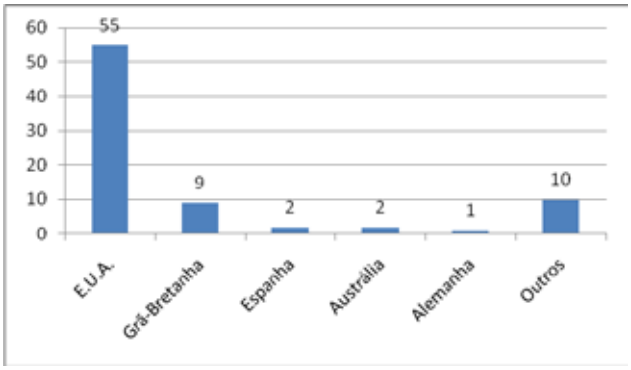

Fig. 10. Localização dos endereços IP compartilhando 5 ou mais arquivos.

VI. Conclusão

Conforme o observado nos testes realizados, a ferramenta EspiaMule identificou uma quantidade muito maior de endereços IP compartilhando arquivos relacionados à exploração sexual infanto-juvenil. Além disso, uma parte significativa desses endereços IP estava localizada no Brasil. Por sua vez, a ferramenta Peer Spectre, integrante do pacote WTK, não identificou nenhum endereço IP localizado no Brasil compartilhando arquivos relacionados à exploração sexual infanto-juvenil. Tal situação é condizente com os dados expostos na Seção I que apontam que a maior parte dos usuários residenciais de Internet no Brasil utiliza o aplicativo eMule. Ademais, cada ferramenta estudada utiliza bases de dados (de códigos *hash* de arquivos) diferentes, podendo influenciar nos resultados obtidos.

Cabe ressaltar que as ferramentas EspiaMule e WTK trabalham com diferentes redes de compartilhamento de arquivos do tipo peer-to-peer e podem ser utilizadas de forma complementar.

Por fim, é importante notar que são necessárias investigações adicionais para determinar a autoria dos resultados obtidos por tais ferramentas.

Referências

[1] DALPIAN, Guilherme M. e BENITES, Carlos A. A., *Ferramenta para monitoramento de redes P2P – Espiamule*. The Second International Conference of Forensic Computer Science, vol 2, n. 1, 2007, p 70-72.

[2] WYOMING DCI ICAC, *Wyoming Toolkit: Instalation and Usage Manual*, version 3, november, 2008.

[3] BARKAY, D., "Peer-to-Peer Computing: Technologies for Sharing and Collaborating on the Net". Intel Press, Agosto de 2001.

[4] GE, Z., FIGUEIREDO, D., JAISWAL, S., KUROSE, J. et Towsley, D., "Modeling Peer-Peer File Sharing Systems", IEEE INFOCOM 2003, São Francisco, CA, USA, Março de 2003.

[5] GUMMADI, K., DUNN, R., SAROIU, S., GRIBBLE, S., LEVY, H., ZAHORJAN, J., "Measurement, Modeling, and Analysis of a Peer-to-Peer File-Sharing Workload", 19th ACM Symposium on Operating Systems Principles (SOSP-19), Bolton Landing, NY, USA, Outubro de 2003.

[6] SCHOLLMEIR, R., "Peer-to-Peer Networking. Applications for and Impacts on Future IP-Based Networks". 3. ITG Fachtagung Netze und Anwengungen, Duisburg, Germany, Março de 2002.

[7] Projeto eMule, http://www.emule-project.net/.

[8] Anonymous – P2P, http://www.anonymous-p2p.org/.

[9] CALAZANS, José. Pesquisa IBOPE/Netrating sobre navegação na Internet em setembro de 2008. Mensagem pessoal enviada para o autor em 22 de junho de 2009.