DOI: 10.5769/C2015014 or http://dx.doi.org/10.5769/C2015014

Catálogo de Fraudes da RNP: 7 anos de experiência no tratamento de fraudes eletrônicas brasileiras

Italo Brito, José Lucas Borges, Lucas Ayres, Paula Tavares, Rogerio Bastos [1]
Edilson Lima, Liliana V. Solha [2]

***Resumo*— As fraudes eletrônicas disseminadas na Internet tornaram-se uma ameaça constante para a população em geral. A cada dia novas e mais sofisticadas técnicas de fraudes são empregadas, levando usuários menos preparados ou atentos à serem vítimas desses ataques. Criado em 2008, o Catálogo de Fraudes da Rede Nacional de Ensino e Pesquisa (RNP), consolida-se como um importante repositório de fraudes eletrônicas brasileiras disseminadas por e-mail. Este artigo apresenta o funcionamento do Catálogo de Fraudes da RNP, estatísticas e tendências observadas, além de oportunidades de trabalho que podem ser desenvolvidos para melhorar a segurança dos usuários de Internet brasileiros.**

***Palavras-Chave*— Fraudes, Catálogo de Fraudes, fraude eletrônica, e-mail, *phishing*, *ICCyber*.**

***Abstract*— The electronic frauds all over the Internet have become a recurring threat to all the people. Everyday, newer and more sophisticated fraud techniques are deployed, causing the less prepared or the less alerted users to be victims of those attacks. The Frauds Catalog of the Brazilian Academic and Research Network (RNP), created in 2008, consolidates itself as an important repository of brazilian electronic frauds disseminated through e-mail. This paper presents RNP Frauds Catalog, how it works, observed statistics and trends, and future work opportunities that can improve the security of Brazilian Internet users.**

***Keywords*— Frauds, frauds catalog, eletronic fraud, e-mail, *phishing*, *ICCyber*.**

I. Introdução

Comunicação sempre foi o principal fator no estabelecimento de relações. Diversas formas de comunicação fáceis e de grande agilidade apresentam-se nos dias atuais e o e-mail continua sendo um dos principais meios de comunicação digital. Apesar de ter sofrido uma desaceleração nos últimos anos, o número de contas de e-mail continua aumentando, bem como o volume de mensagens transitadas. Atribui-se como motivo desse contínuo crescimento, a facilidade de uso e a maior formalidade e confiança atribuída à mensagem de e-mail, se comparada a outros meios de comunicaçao. Em 2014, segundo resultados apresentados pelo grupo Radicati [1], cerca de 4,1 bilhões de contas de e-mails sendo utilizadas em todo mundo e 108,7 bilhões de mensagens trafegando por dia.

Embora seja um serviço de comunicação utilizado em larga escala desde o final da década de 90, as ferramentas de e-mail ainda possuem, em sua maioria, uma verificação do conteúdo precária, o que se presta a um possível mal uso. Por esse motivo, o e-mail é bastante utilizado para disseminação de fraudes eletrônicas [2]. As fraudes frequentemente apelam para a inocência, tal como a não validação de uma suposta mensagem de seu banco solicitando dados pessoais; ou curiosidade dos usuários, por exemplo, a possibilidade de ver fotos exclusivas de um evento recentemente noticiado pela mídia.

Os e-mails fraudulentos estão, em sua maioria, relacionados a golpes digitais como o *phishing*. O *phishing* ocorre quando o golpista, utilizando-se de meios digitais e de engenharia social, tenta obter dados pessoais, senhas ou informações financeiras da vítima. O usuário pode ser seduzido de diversas formas, como por exemplo por meio de uma página de compras em promoção, solicitações de atualização ou recadastramentos, que caso não ocorram acarretarão em prejuízo ao usuário, ou ainda como promoções relacionadas a cartões de crédito, a companhias aéreas ou outras envolvendo o preenchimento de formulários, levando o usuários desavisados a clicarem em links falsos criados pelo atacante.

Muitas vezes o link falso dá origem ao *download* de um *malware*, um tipo de programa desenvolvido unicamente para a execuçao de ações maliciosas. Ao ser instalado ou executado no computador da vítima, o *malware* pode iniciar funções tais como envio de spam, roubo de informações confidenciais dos usuários ou até mesmo realizar ataques contra outras máquinas.

O Catálogo de Fraudes da RNP foi criado em 2008, com o objetivo de coletar fraudes recebidas por e-mail pela população em geral e analisar, filtrar e catalogar essas fraudes, criando um repositório de mensagens conhecidamente fraudulentas e alertando a comunidade sobre como se proteger desse tipo de ataque. Criado pelo Centro de Atendimento a Incidentes de Segurança da RNP (CAIS/RNP) e mantido atualmente em parceria com o Ponto de Presença da RNP na Bahia (PoP-BA/RNP), até onde se tem conhecimento, o Catálogo de Fraudes da RNP é a única fonte de informações aberta e online sobre fraudes eletrônicas no Brasil, sendo bastante utilizado pela população em geral como uma base de conhecimento para validação de mensagens de e-mail suspeitas.

Este artigo está estruturado em seis seçoes, sendo esta Seçao da Introduçao a primeira. Na Seção II são apresentados outros trabalhos, todos internacionais, que abordam a análise e registro de fraudes eletrônicas. Já na Seção III, discute-se o processo de tratamento de fraudes, detalhando-se as etapas deste processo, que incluem: recebimento, triagem, interação com a fraude e catálogo. A Seção IV traz algumas estatísticas

[1]Italo Brito, José Lucas Borges, Lucas Ayres, Paula Tavares, Rogerio Bastos. Ponto de Presença da RNP na Bahia- PoP-BA/RNP, Universidade Federal da Bahia - UFBA, Salvador-BA, E-mails: {italo,lucasborges,lucasayres,paulatavares,rogeriobastos}@pop-ba.rnp.br. [2]Edilson Lima, Liliana V. Solha. Centro de Atendimento a Incidentes de Segurança da RNP - CAIS/RNP, Campinas-SP, E-mails: {edilson.lima,liliana.solha}@rnp.br.

e tendências observadas no tratamento das fraudes e a Seção V apresenta alguns benefícios que o Catálogo de Fraudes proporciona para a comunidade em geral. Por fim, a Seção VI conclui e apresenta trabalhos futuros para esta iniciativa.

II. Trabalhos Relacionados

Os trabalhos anteriores na área de fraudes eletrônicas (*phishing*) focam em mecanismos de detecção automática de fraudes, técnicas utilizadas pelos atacantes e relatórios de atividades envolvendo *phishing*.

Detecção automática é um assunto bastante abordado em artigos dessa área. Basnet et al. em [7] tentou criar um sistema de detecção automática baseado em regras. Estas regras levam em consideração diversas características, como IPs nas URLs, palavras-chave comuns a sites fraudulentos, número de caracteres especiais na URL, envio de formulários utilizando TLS/SSL, presença na blacklist do Google Safe Browsing [5], número de pontos e tamanho da URL, entre outras. Fette et al. em [8] propõe um método de classificação de e-mails baseado no potencial que cada e-mail tem de ser isca para um ataque de *phishing*. Seu algoritmo de detecção leva em conta características únicas identificadas nos emails de *phishing* e a saída de um filtro de spam, que eles identificaram ser bastante eficiente para este tipo de detecção.

McGrath et al. em [9] examina o modus operandi dos atacantes através da anatomia das URLs e domínios dos sites fraudulentos, do registro e tempo de ativação dos domínios de *phishing*, e das máquinas utilizadas para hospedar tais sites. Os resultados podem ser utilizados como heurísticas na filtragem de e-mails de *phishing* e na identificação de registros de domínios suspeitos. Garera et al. em [10] estuda as URLs utilizadas em ataques de *phishing* e tenta descobrir se elas realmente pertencem a um ataque de *phishing*, sem utilizar qualquer conhecimento da página em si. Ele identificou algumas técnicas utilizadas pelos *phishers* para enganar as vítimas, como mascarar o *host* com um endereço de IP, mascarar o *host* com outro domínio, criar domínios similares ao de organizações conhecidas e criar URLs muito grandes para confundir a vítima.

A RSA [11] produz relatórios mensais da quantidade de ataques envolvendo *phishing* e suas principais características. O Anti-Phishing Working Group (APWG) publica relatórios trimestralmente [12] com as tendências de ataques de *phishing*, informando a quantidade de marcas utilizadas como alvo pelos *phishers*, os setores mais atacados da indústria, os países que mais hospedam sites de *phishing* e cavalos-de-tróia utilizados nestes ataques, e a distribuição dos *phishings* por TLD (top-level domain).

III. Processo de tratamento

Desde 2008, o Centro de Atendimento a Incidentes de Segurança da RNP (CAIS/RNP) mantém uma base de dados de fraudes eletrônicas encaminhadas por usuários da Internet. Todos os e-mails recebidos pelo CAIS apontados como alerta de uma potencial fraude são analisados e catalogados em uma ferramenta web, disponibilizando as informações coletadas como fonte de consulta através do site da iniciativa[1]. O objetivo do catálogo é, portanto, apoiar a comunidade brasileira na identificação e conscientização sobre os principais golpes eletrônicos que estão sendo veiculados na Internet.

A partir dos e-mails recebidos pelo CAIS, é realizada uma triagem inicial. Atualmente, cerca de 15.000 mensagens são tratadas a cada mês, desse total são descartados os spams e as mensagens em língua estrangeira, posteriormente, descarta-se as mensagens repetidas. Com isso, são catalogadas uma média de 200 novas fraudes por mês. Após a mensagem ser classificada como fraude, inicia-se o processo de identificação das principais características do e-mail, tais como o uso de redirecionamento para sites falsos e/ou a presença de arquivos maliciosos em anexo ou disponíveis para *download*.

São registrados, no Catálogo de Fraudes, o corpo do e-mail na forma de texto e imagem através de captura de tela. As mensagens que direcionam o usuário para sites fraudulentos, também têm as páginas do site registradas como imagem, para isso é feito uma interação com esses sites, a fim de coletar o maior número de informações. Quando há um *malware* anexado ou disponível para *download*, é utilizada a ferramenta VirusTotal [3] para análise do arquivo malicioso, sendo registrado o nome do arquivo que está relacionado a sua ação em uma máquina, como por exemplo "trojan" que dá acesso a usuários maliciosos à máquina da vítima e *malware* contendo a palavra "Win" geralmente são direcionados a máquinas com sistema operacional Windows. Estes dados são utilizados para melhor informar o usuário.

As informações, tais como assunto da mensagem, tipo, classificação, nome do *malware*, hash md5 do *malware* são registradas no catálogo juntamente com as imagens. As principais tarefas relacionadas ao Catálogo de Fraudes são atualmente executadas manualmente, no entanto, objetivando aumentar a eficiência e os resultados obtidos, encontram-se em desenvolvimento novas ferramentas para a automatização de algumas etapas deste processo. Estas melhorias visam tornar este processo mais eficaz e possibilitar a identificaçao de um número maior de fraudes.

IV. Experiências obtidas na manutenção do Catálogo de Fraudes

Ao longo desses sete anos de tratamento e análise de fraudes, foi possível observar diversas tendências no comportamento dos e-mails fraudulentos. O entendimento dessas tendências oferece condições de desenvolver guias de recomendações e ferramentas automatizadas que possam ajudar no combate à disseminação dessas fraudes. Essa seção apresenta algumas estatísticas, tais como: quantidade de e-mails recebidos, número de fraudes analisadas, número de fraudes por categoria, etc. Assim também, são apresentados resultados de análises de URLs maliciosas e tendências observadas no Catálogo de Fraudes da RNP.

A. Estatísticas de fraudes catalogadas

Os e-mails recebidos passam por uma triagem e são classificados como: spam, fraude ou mensagem em língua estrangeira.

[1] http://www.rnp.br/servicos/seguranca/catalogo-fraudes

Grande parte dos e-mails encaminhados por usuários são mensagens de spam. A figura 1 contrasta a quantidade total de e-mails recebidos com a quantidade de mensagens classificadas como fraudes ou mensagens em língua estrangeira, no período de abril de 2014 a abril de 2015.

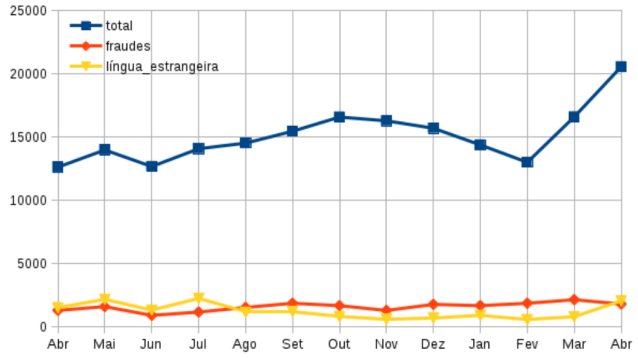

Fig. 1. E-mails tratados em 2014/Abr - 2015/Abr

Com base nos assuntos abordados nas mensagens classificadas como fraude, estas são categorizadas da seguinte maneira: bancos, documentos, cartões de crédito, e-commerce, empresas aéreas, governo, redes sociais e internet, seguradoras, serviços de pagamento, e outras.

O gráfico da figura 2 mostra o acumulado de fraudes por categoria, no período de abril de 2014 a abril de 2015. As fraudes mais frequentes exploram assuntos relacionados às instituições bancárias, reunidas na categoria bancos, e a transações financeiras, tais como pagamento de boletos e faturas, cheques, comprovantes de depósitos e transferências, orçamentos, dentre outros, reunidos na categoria documentos.

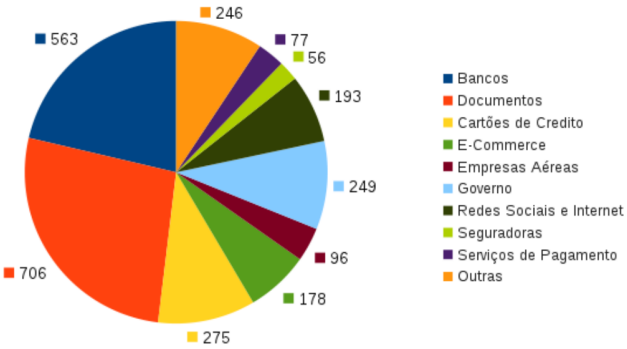

Fig. 2. Categorização das fraudes catalogadas em 2014/Abr - 2015/Abr

Também se destacam-se as fraudes relacionadas a empresas de cartões de crédito e fraudes que utilizam o nome de instituições governamentais. É comum temas como intimações judiciais, serviços postais, cadastro em promoções de empresas de cartões de crédito, dentre outros.

B. Análise de URLs maliciosas

Analisando-se os e-mails de fraudes, constata-se que cerca de 90% contêm URLs para sites ou arquivos maliciosos no corpo da mensagem. Diante dessa informação, verificou-se que o bloqueio das URLs maliciosas é um importante mecanismo de proteção para os usuários.

Atualmente, há diversos serviços de reputação de URLs. O Google Safebrowsing [5] é o serviço do Google para verificação de URLs maliciosas. Navegadores web como o Firefox, Google Chrome e Safari utilizam esse serviço para alertar os usuários quando as mesmas são acessadas.

O Phishtank [4] é uma comunidade baseada no serviço anti-phishing, usado por empresas como Opera, WOT, Yahoo Mail para verificar se uma URL é considerada *phishing*. A Microsoft oferece o filtro SmartScreen [6] para os aplicativos do seu sistema operacional, contudo não disponibiliza uma API de consulta pública.

Utilizando as APIs de consulta dos serviços Google Safebrowsing e Phishtank realizamos uma análise das URLs encontradas nas fraudes catalogadas. As figuras 3 e 4 mostram a quantidade de URLs analisadas e a quantidade de URLs detectadas como maliciosas nos serviços Google Safebrowsing e Phishtank.

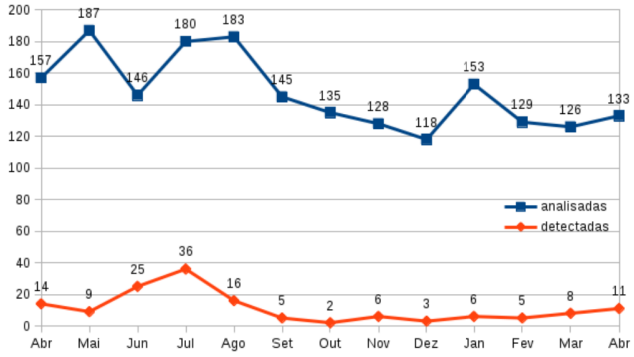

Fig. 3. Análise de URLs no Google Safebrowsing

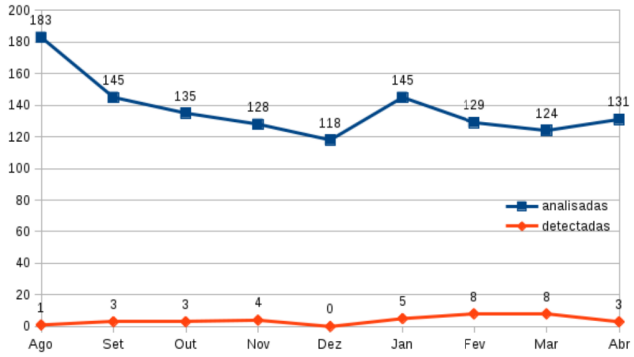

Fig. 4. Análise de URLs no Phishtank

Como pode ser observado nos gráficos apresentados, os serviços de reputação avaliados são pouco eficientes para as URLs utilizadas nas fraudes destinadas aos usuários brasileiros. Isso mostra a necessidade de um serviço direcionado a estes usuários. Diante dessa demanda, a RNP iniciou o desenvolvimento de um serviço de verificação de URLs maliciosas voltado para a comunidade brasileira.

C. Principais fraudes

Similar às campanhas de spam, as fraudes eletrônicas também apresentam campanhas direcionadas a determinados acontecimentos ao longo do ano. No ano de 2014 e inclusive em 2015, puderam ser identificadas algumas fraudes que se destacaram por abordarem assuntos relacionados a acontecimentos de relevância nacional.

No período da Copa do Mundo este tema foi largamente abordado relacionado a promoções e compra de ingressos que redirecionava o usuário para páginas falsas que requeriam seus dados pessoais e financeiros. Na época das eleições também foi grande a quantidade de *phishings* abordando temas como cadastramento biométrico, regularização do título eleitoral, falsas notícias sobre os candidatos à presidência, processo de seleção de mesários para zonas eleitorais e listas a favor de um Impeachment, fraudes que, em sua maioria continham arquivos maliciosos utilizados para infectar a máquina do usuário. O ENEM foi utilizado como tema em diversas fraudes que requeriam o recadastramento devido a irregularidades na inscrição, porém, a URL citada na suposta mensagem do ENEM levava o usuário a executar inadivertidamente um *malware*. Assuntos como imposto de renda deram origem a fraudes contendo falsos documentos.

Algumas campanhas fraudulentas apresentam-se regularmente durante todo o ano. Fraudes relacionadas a bancos são enviadas cotidianamente solicitando atualizações e recadastramentos para que supostamente não ocorra bloqueio da conta ou indisponibilização de serviços. Fraudes associadas a comprovantes, boletos, notas fiscais, faturas, pedidos de orçamento estão relacionadas em sua maioria com arquivos maliciosos. Essas fraudes que se repetem ao longo do ano possuem maiores chances de ludibriar a vítima por retratar assuntos comuns, de forma simples e de interesse pessoal.

V. Benefícios do catálogo

O Catálogo de fraudes é uma importante contribuição da RNP tanto para a comunidade acadêmica como também à comunidade brasileira em geral. As fraudes ali catalogadas, por serem rigorosamente analisadas, são uma fonte confiável de informações sobre fraudes eletrônicas brasileiras. Os usuários, em geral, podem fazer uso dessas informações para consultar por período, identificar campanhas de fraudes, e também comparar o texto e imagem das mensagens recebidas com fraudes conhecidas.

Como consequência do trabalho com as fraudes eletrônicas, alguns guias de boas práticas, respostas para perguntas frequentes e cartilhas de segurança foram desenvolvidas e disponibilizadas no site da iniciativa[2].

Já para a comunidade de Segurança da Informação, esse catálogo pode ser usado para maior entendimento das fraudes direcionadas ao público brasileiro, uma vez que a maioria dos trabalhos anteriores apresenta dados de fraudes internacionais. Nota-se, inclusive, a carência por repositórios dessa natureza aqui no Brasil, de forma que possam ser usados por ferramentas de segurança automatizadas para evitar que os usuários se tornem vítimas (ex: filtros de conteúdo, *plugins* de navegadores web, etc). Neste sentido, é interesse da RNP incorporar ao Catálogo de Fraudes o registro das URLs utilizadas nas fraudes, agregando um mecanismo de reputação às URLs de forma que possam ser usadas em ferramentas de clientes - este trabalho atualmente encontra-se em andamento. Por outro lado, a análise dessas URLs pode desvendar padrões, comumente utilizados pelos atacantes, para iludir seus usuários, comparando-os com características das fraudes identificadas ao redor do mundo.

[2]http://www.rnp.br/servicos/seguranca/educacao-e-conscientizacao-seguranca

VI. Conclusões

O crescimento na disseminação de fraudes eletrônicas, via e-mail, redes sociais e outras mídas eletrônicas, associado com a falta de conhecimento e até discernimento de muitos usuários na Internet, evidenciam a necessidade e a importância de uma base de conhecimento das fraudes eletrônicas, que possa ser usada não apenas como uma fonte de consulta e validação de mensagens desconhecidas, mas também que possa ser usada pelas organizações na implantação de filtros e outros mecanismos de segurança a fim de evitar que os usuários sejam vítimas desse tipo de ataque.

Este artigo apresentou o processo de análise, triagem e registro do Catálogo de Fraudes da RNP, bem como algumas estatísticas e tendências observadas nesse processo. É notório observar o crescimento na quantidade das fraudes e também o nível de sofisticação nas páginas de captura de informações dos usuários. Dessa maneira é importante que novos mecanismos de filtro automatizado de sites e mensagens fraudulentas sejam implantados, a fim de minimizar a quantidade de usuários que estão sujeitos a esses golpes. Apresentou-se também uma análise de ambientes de filtros de URL, a saber, o projeto *Google Safe browsing* e *Phishtank*, onde evidênciou-se a necessidade de uma base de dados voltada para a realidade brasileira de fraudes eletrônicas, haja visto que a grande maioria das fraudes não é encontrada nessas bases de dados internacionais (7.6% de detecção pelo *Safe browsing* e 2.82% de detecção pelo *Phishtank*).

Como trabalhos futuros, espera-se aprimorar a apresentação do Catálogo de Fraudes para os usuários, adicionando novos mecanismos de busca no site, gráficos, e formulários de validação de e-mail; implantar um catálogo de URLs maliciosas observadas nacionalmente, fornecendo uma API de consulta para viabilizar filtros automatizados nas instituições (trabalho em andamento), investigar e propor mecanismos de detecção automatizada de fraudes, para que o processo de triagem do Catálogo de Fraudes possa ser automatizado e assim a quantidade de fraudes analisadas seja maior.

Referências

[1] The Radicati Group, Inc. Email Statisctics Report, 2014-2018. `http://www.radicati.com/?p=10644`, acessado em 12/05/2015.

[2] Ollmann, Gunter. *The Phishing Guide: Understanding & Preventing Phishing Attacks.* IBM Int. Sec. Sys. Disponível em `http://www-935.ibm.com/services/us/iss/pdf/phishing-guide-wp.pdf`, acessado em 29/04/2015.

[3] VirusTotal. `https://www.virustotal.com/`, acessado em 29/04/2015.

[4] Phishtank. `https://www.phishtank.com/`, acessado em 29/04/2015.

[5] Google's Safe Browsing Diagnostic Tool. https://developers.google.com/safe-browsing/, acessado em 04/05/2015.
[6] Filtro SmartScreen. http://windows.microsoft.com/pt-br/internet-explorer/products/ie-9/features/smartscreen-filter, acessado em 02/05/2015.
[7] Basnet, Ram B., Andrew H. Sung, and Quingzhong Liu. *Rule-based phishing attack detection.* International Conference on Security and Management (SAM 2011), Las Vegas, NV. 2011.
[8] Fette, Ian, Norman Sadeh, and Anthony Tomasic. *Learning to detect phishing emails.* Proceedings of the 16th international conference on World Wide Web. ACM, 2007.
[9] McGrath, D. Kevin, and Minaxi Gupta. *Behind Phishing: An Examination of Phisher Modi Operandi.* LEET 8 (2008): 4.
[10] Garera, Sujata, et al. *A framework for detection and measurement of phishing attacks.* Proceedings of the 2007 ACM workshop on Recurring malcode. ACM, 2007.
[11] RSA Online Fraud Resource Center. http://www.emc.com/emc-plus/rsa-thought-leadership/online-fraud/index.htm#!resources, acessado em 29/04/2015.
[12] APWG Reports. http://www.antiphishing.org/resources/apwg-reports/, acessado em 29/04/2015.