SpamBands: a Methodology to Identify Sources of Spam Acting in Concert

E. Fazzion, P. H. B. Las-Casas, O. Fonseca, D. Guedes, W. Meira Jr, C. Hoepers, K. Steding-Jessen and M. H. P. Chaves

***Abstract*— In 2012, estimates indicated that 68.8% of all e-mail traffic was spam, what suggests this is still a relevant problem. Recently, some works have focused on the analysis of spam's traffic inside the network, analyzing the protocols used and the AS which originate the traffic. However, those works usually do not consider the relationships between the machines used to send spam. Such an analysis could reveal how different machines may be used by a single spammer to spread his messages, helping us to understand their behavior. To that end, this work proposes a methodology to cluster the machines used by spammers based on the concept of spam campaigns. The groups identified were characterized to identify different aspects of the spam dissemination process, which suggest different orchestration strategies being used.**

***Keywords*— SpamBands, Spam traffic, Spam orchestration.**

I. Introdução

Há muitos conceitos sobre o que é *spam*, porém todos têm uma base comum: um *spam* é uma mensagem de email de caráter não individual e não solicitada, que é disseminada em larga escala pela rede. As motivações daqueles que realizam essa prática, os *spammers*, são diversas, sendo as mais comuns a venda de produtos, a disseminação de *malware* e ataques de *phishing* [1]. Segundo a companhia Pingdom, cerca de 144 bilhões de mensagens de email foram enviados por dia, em 2012, sendo 68,8% delas *spam* [2]. Isso mostra que recursos para enviar e armazenar 99 bilhões de mensagens, por dia, foram desperdiçados, o que leva a sérios prejuízos financeiros, como revelado em outros trabalhos [3]. Além disto, existe um prejuízo social, onde mensagens legítimas são perdidas por má classificação de filtros de *spam* ou por excesso de tráfego ocasionado por grandes volumes de *spam*[4].

Existem diversas facetas consideradas no combate ao *spam*. Muitos estudos buscam entender o problema do ponto de vista do destinatário e auxiliar na construção de filtros eficazes que descartem mensagens indesejáveis. Outros fazem a análise do comportamento do *spammer* na rede, para entender como o *spam* é disseminado, de onde ele se origina e como ele atravessa a rede sem que os transmissores sejam facilmente identificados. O objetivo, nesse caso, é identificar comportamentos na rede que permitam bloquear as mensagens antes que elas atravessem a rede e consumam recursos para sua filtragem e possível armazenamento [5].

Em ambos os casos, fica visível que o combate ao *spam* requer o entendimento de um sistema complexo de ofuscação usado pelo *spammer* em sua atividade. Esse sistema exige uma complexa orquestração de atores e recursos, cuja existência é reconhecida mas que normalmente é invisível para o profissional que se dedica a esse combate. Para se manter oculto, o *spammer* busca disfarçar sua localização na rede, seja enviando suas mensagens a partir de múltiplas origens, como máquinas infectadas que se organizam em *botnets*, ou usando servidores especializados que podem por sua vez se aproveitar de máquinas mal-configuradas na rede para se ocultar dos destinatários. Além disso, *spammers* também utilizam programas de transmissão que geram diversas mensagens diferentes como versões de um mesmo conteúdo básico, a fim de tentar ludibriar os filtros baseados em conteúdo [4]. Nesse processo, tem importância o conceito de *campanhas de spam*, que são grupos de mensagens que possuem um mesmo objetivo, mas que foram alteradas por métodos de ofuscação para tentar ludibriar filtros [6].

Este trabalho utiliza uma abordagem que combina aspectos de campanhas com aspectos de comportamento de rede a fim de tentar lançar mais luz sobre esse elemento orquestrador subjacente ao processo de envio de *spam*. Para este fim, utilizamos tanto elementos baseados no conteúdo da mensagem, para permitir a identificação das *campanhas de spam*, quanto elementos do tráfego de rede, para identificar as máquinas originadoras de cada campanha. Com isso, propomos um método capaz de identificar os grupos de máquinas na rede que se encontram em um certo momento sob o controle de um orquestrador oculto, o *spammer*. A esses grupos denominamos *SpamBands*.

Segundo a abordagem adotada neste trabalho, um(a) *SpamBand* é um grupo de máquinas correlacionadas pelo fato de terem enviado mensagens identificadas como pertencentes a um mesmo conjunto de campanhas de *spam*. Utilizando essa estrutura em nossas avaliações, conseguimos mostrar relações importantes como o período de atividade de cada *SpamBand* e a forma como o *spammer* escolhe o protocolo utilizado. Com relação ao período de atividade, mostramos a tendência desses grupos se manterem estáveis ao longo do tempo, podendo

E. Fazzion, Universidade Federal de Minas Gerais (UFMG), Belo Horizonte, MG, Brasil, elverton@dcc.ufmg.br

P. H. B. Las-Casas, Universidade Federal de Minas Gerais (UFMG), Belo Horizonte, MG, Brasil, pedro.lascasas@dcc.ufmg.br

O. Fonseca, Universidade Federal de Minas Gerais (UFMG), Belo Horizonte, MG, Brasil, osvaldo.morais@dcc.ufmg.br

D. Guedes, Universidade Federal de Minas Gerais (UFMG), Belo Horizonte, MG, Brasil, dorgival@dcc.ufmg.br

W. Meira Jr, Universidade Federal de Minas Gerais (UFMG), Belo Horizonte, MG, Brasil, meira@dcc.ufmg.br

C. Hoepers, Centro de Estudos para Resposta e Tratamento de Incidentes de Segurança no Brasil (CERT.br/NIC.br), São Paulo, SP, Brasil, cristine@cert.br

K. Steding-Jessen, Centro de Estudos para Resposta e Tratamento de Incidentes de Segurança no Brasil (CERT.br/NIC.br), São Paulo, SP, Brasil, jessen@cert.br

M. H. P. Chaves, Centro de Estudos para Resposta e Tratamento de Incidentes de Segurança no Brasil (CERT.br/NIC.br), São Paulo, SP, Brasil, mhp@cert.br

se estender por diversas campanhas e que a técnica pode identificar, como efeito adicional, possíveis partes de redes *botnets*. Quando consideramos a forma como as mensagens são enviadas, observamos que, *em geral*, *SpamBands* utilizam apenas *proxies* (HTTP ou SOCKS) ou apenas *mail relays* abertos (SMTP) em seus envios, apesar de algumas *SpamBands* apresentarem um comportamento híbrido, utilizando os dois tipos de protocolos.

A definição de *SpamBand* pode facilitar a identificação de *botnets* e outras infra-estruturas de distribuição utilizadas pelos *spammers*. Com isso, ações podem ser desenvolvidas para impedir a ação das máquinas envolvidas, removendo-as da rede ou procedendo à remoção de qualquer *malware* nelas instalado. Além disso, pela identificação dos grupos pode se tornar mais eficaz o uso de *blacklists* no bloqueio ao *spam*: se uma máquina é identificada como fazendo parte de um grupo que contém elementos já incluídos em uma lista negra, essa nova máquina também pode ser automaticamente adicionada àquela lista.

II. Metodologia de identificação de *SpamBands*

O conceito de *SpamBands* foi desenvolvido durante a análise dos dados de *spam* coletados em diversos pontos da Internet, onde percebemos que várias origens surgiam na análise do spam observado em diferentes pontos da rede. Nesta seção detalhamos a metodologia proposta para a identificação das *SpamBands* e um exemplo real de aplicação que ressalta alguns elementos importantes da proposta.

Como mencionado, a base do conceito de *SpamBands* é a premissa de que máquinas que enviam mensagens pertencentes às mesmas campanhas são controladas por um mesmo agente orquestrador, estando, assim, relacionadas a uma mesma origem. A relação entre máquinas e campanhas pode ser modelada como um grafo G, onde as máquinas são vértices e há uma aresta entre duas máquinas se elas enviaram mensagens associadas a uma mesma campanha. A Fig. 1 ilustra a construção desse grafo.

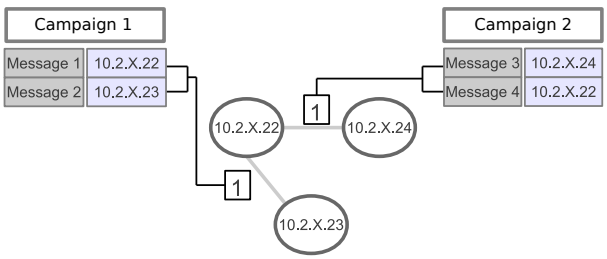


Figura 1. Modelo de grafo para relação entre campanhas e endereços IP.

A partir do grafo G, um *SpamBand* pode ser identificado como um sub-grafo denso (diversas origens que compartilham um mesmo conjunto de campanhas). A identificação desses subgrafos pode ser obtida aplicando-se algoritmos de agrupamento de grafos; entretanto, tais algoritmos tendem a ser bastante complexos e difíceis de calibrar [7]. Com base nas características particulares do problema em questão, adotamos uma estratégia mais simples e interativa, descrita a seguir.

Inicialmente, cada componente conectado de G poderia ser identificado como um *SpamBand*. Entretanto, aspectos práticos exigem que essa definição seja refinada. Por exemplo, quando um endereço IP pode se referir a diferentes máquinas atrás de um mecanismo de NAT: duas máquinas podem estar atuando sob coordenações diferentes, mas serem vistas no resto da rede como um mesmo endereço de origem. Em outros casos, um endereço é visto participando de uma campanha até certo instante do dia e a partir de então passa a participar de outra. Os nós referentes a esses endereços IP aparecem no grafo como nós de ligação entre sub-grafos mais densos, que na prática se referem a *SpamBands* diferentes.

A forma adotada para identificar esses casos e isolar os *SpamBands* envolvidos foi utilizando-se o conceito de *betweenness*, que mede o grau de centralidade de nós em um grafo. Essa métrica quantifica o número de caminhos mínimos entre todos os pares de nós no grafo que passam por um vértice em questão. A premissa é que, se alguns vértices possuem um valor de *betweenness* muito elevado em relação ao que seria esperado para um grafo fortemente conectado, existe uma chance maior desses vértices conectarem dois sub-grafos internamente mais densos. Assim, se removemos esses vértices, acentuamos a separação entre os sub-grafos densos desejados.

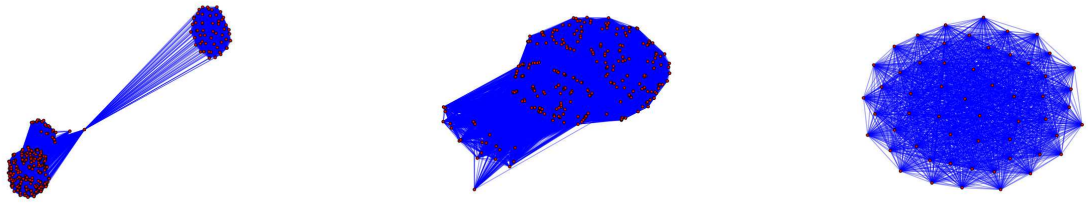
A determinação de *SpamBands* é então apresentada no algoritmo 1, que recebe três parâmetros de entrada: o grafo (**G**), o limiar de *betweenness* mínimo a ser considerado (**limiar_bt**) e o número máximo de endereços IP (vértices) que podem ser removidos para dividir um componente (**limiar_ips**). O primeiro passo determina os componentes conectados de G, que constituem uma primeira aproximação dos *SpamBands*. A seguir, identificamos sub-grafos densos em cada componente conectado removendo nós com *betweenness* acima de **limiar_bt**, respeitando o limite **limiar_ips**, que define o tamanho mínimo de um sub-grafo denso, para evitar a geração de conjuntos muito pequenos. O algoritmo retorna o conjunto S que contém todos os *SpamBands*.

Por exemplo, a Fig. 2(a) mostra um dos componentes conectados com maior número de máquinas observados em um dos dias da nossa análise. Claramente, podemos verificar que há pelo menos dois grupos praticamente disjuntos de nós, unidos por um nó que aparece entre eles. Aplicando o algoritmo 1 naquele componente conectado, isolamos os dois *SpamBands* relativos aos grupos mais densos, mostrado nas figuras 2(b) e 2(c).

Analisamos os *SpamBands* revelados através do componente conectado da Fig. 2(a). O *SpamBand* da Fig. 2(b) está distribuído em quatro ASes (17816, 17623, 4837 e 17430). Por outro lado, apesar do *SpamBand* da Fig. 2(c) estar localizado no mesmo *Country Code*, seu único AS (4134) difere de todos os outros ASes do *SpamBand* da Fig. 2(b), o que traz uma forte diferença e sugere que esses *SpamBands* identificam *botnets* distintas.

III. Coleta de dados

Os dados utilizados na análise foram coletados utilizando-se oito *honeypots* de baixa interatividade instalados em diferentes



(a) *SpamBand* original (b) *SpamBand* 1 (c) *SpamBand* 2

Figura 2. *SpamBands*: componente original e aqueles revelados a partir da aplicação do algoritmo 1.

Algorithm 1: *SpamBands* (Grafo G, Real limiar_bt, Real limiar_ips)

```
S = ∅;
C=G.ComponentesConectados(); ;
for comp em C do
    ips_a_remover = ∅ ;
    for ip em comp do
        if ip.Betweenness() >
          limiar_bt*comp.MaiorBetweenness() then
            ips_a_remover.Adiciona(ip);
        end
    end
    if ips_a_remover.Tamanho() >
      limiar_ips*comp.Numvertices() then
        S += comp;
    end
    else
        S += comp.RemoveVertices(ips_a_remover);
    end
end
retorna S;
```

partes do mundo: Austrália (AU-01), Áustria (AT-01), Brasil (BR-01 e BR-02), Equador (EC-01), Holanda (NL-01), Taiwan (TW-01) e Uruguai (UY-01). A distribuição desses *honeypots* teve por objetivo capturar dados de diferentes pontos da Internet, a fim de obter uma visão mais global do spam que viaja pela rede.

Todos os *honeypots* foram desenvolvidos de modo a simular computadores com *proxies* HTTP e SOCKS e *mail relays* SMTP abertos, que frequentemente são abusados para o envio de spam. Quando uma máquina se conecta à porta 25 de um dos *honeypots*, ela tem a impressão de estar interagindo com um servidor SMTP operando como um *open relay*, que repassa mensagens de correio para outros servidores. Já uma máquina que se conecta a um *honeypot* através dos protocolos HTTP ou SOCKS, é levada a crer que é capaz de estabelecer conexões para outros servidores SMTP na rede. Toda a interação do atacante com o suposto servidor de correio é registrada e as mensagens de spam são armazenadas localmente — nenhuma mensagem de spam é realmente entregue ao seu destino, exceto mensagens classificadas como mensagens de teste, segundo regras pré-definidas. Periodicamente, ao longo de cada dia, todo o spam armazenado nos *honeypots* é copiado para os servidores centrais do projeto.

O período de coleta usado nesta análise foi de 07/10/2013 a 25/10/2013, totalizando 19 dias consecutivos. A Tabela I oferece uma visão geral dos dados coletados.

Cerca de 225 milhões de mensagens foram coletadas, provenientes de endereços IP associados a 93 *country codes* distintos. Apesar do protocolo SOCKS ser o responsável pela maior parte do tráfego, representando 51,8% das mensagens enviadas, o número de endereços IP que utilizam o protocolo SMTP é maior, com 69,4% do total, mesmo enviando um número inferior de mensagens.

A Tabela II mostra o número de IPs, o número de mensagens e o número de ASes observados em cada *honeypots*. É importante notar que existe uma sobreposição de IPs entre *honeypots*, que indica que grupos de disseminação de spam estão atuando em mais de um coletor. Este fato será detalhado posteriormente.

IV. Resultados

Nesta seção, apresentamos os principais resultados obtidos utilizando a técnica descrita na Seção II. Inicialmente, na Subseção IV-A, fazemos um estudo de caso detalhado de forma a mostrar diferentes tipos de *SpamBands* e como estes atuam.

Na Subseção IV-B, damos uma visão geral do comportamento dos *SpamBands* encontrados nos *honeypots* e uma possível orquestração de máquinas. Ainda mais, mostramos, por meio de um exemplo, que existem *SpamBands* atuando em diferentes *honeypots*, reforçando a existência de uma orquestração e a eficácia da técnica.

A Subseção IV-C mostra um resultado imediato, obtido através do estudo dos *SpamBands*, no aprimoramento de *blacklists*. Por último, na Seção IV-D, apresentamos um estudo temporal dos *SpamBands* com resultados interessantes, realçando o quão valiosa a técnica exposta neste artigo pode ser no estudo dos *spammers* nessa dimensão.

A. Estudo de caso

Nesta seção, detalhamos os *SpamBands* descobertos nos dados do exemplo ao final da Seção II. Todos os 7 *SpamBands* podem ser vistos na Tabela III.

TABELA I
VISÃO GERAL DA BASE

	HTTP(%)	SMTP (%)	SOCKS (%)	Total
Mensagens (milhões)	76,25 (33,7)	32,82 (14,5)	116,58 (51,8)	225,66
Endereços IP	11135 (29,3)	26313 (69,4)	4372 (11,5)	37895
Prefixos de rede	40 (1,5)	2218 (87,7)	342 (13,5)	2529
Sistemas Autônomos (AS)	11 (1,6)	591 (89,0)	125 (18,8)	664
***Country Codes* (CC)**	6 (6,4)	92 (98,9)	31 (33,3)	93
Volume de Tráfego (GB)	211,18 (28,6)	160,74 (21,7)	365,97 (49,7)	737,90

TABELA II
MENSAGENS E IPS POR *honeypot*

	AT-01	AU-01	BR-01	BR-02	EC-01	NL-01	TW-01	UY-01
Mensagens (milhões)	25,27	6,51	13,89	38,64	16,57	57,52	53,92	13,33
Endereços IP	10438	19420	26762	11261	25494	11053	11145	10138
ASes	330	330	473	142	274	130	122	327

TABELA III
SpamBands DESCOBERTOS NO *honeypot BR-01* DO EXEMPLO AO FINAL DA SEÇÃO II

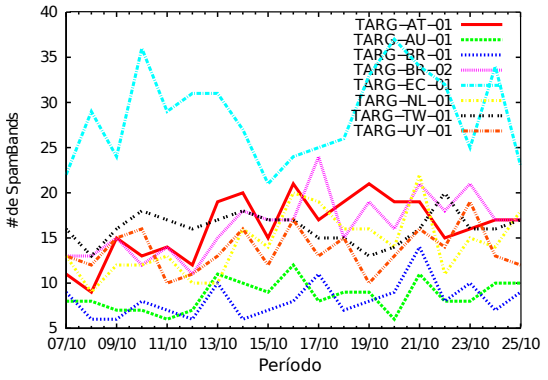
	Msg	IPs	ASes	CC (Top)	SMTP (%)	SOCKS (%)	HTTP (%)	XBL	PBL	Número de horas ativo
***SpamBand* 1**	48.244	971	1	1 (TW)	100	0	0	55	971	24
***SpamBand* 2**	475.971	910	198	52 (CN)	100	0	0	636	597	24
***SpamBand* 3**	2.711	303	4	1 (CN)	100	0	0	224	257	18
***SpamBand* 4**	1.795	56	1	1 (CN)	0	100	0	1	53	23
***SpamBand* 5**	35.389	200	96	26 (BR)	0	100	0	0	56	24
***SpamBand* 6**	28.680	5	1	1 (TW)	0	100	0	0	5	24
***SpamBand* 7**	16.679	3	1	1 (TW)	0	100	0	0	3	24

Entre os *SpamBands* da tabela, podemos identificar três grupos. O primeiro é composto pelos *SpamBands* 6 e 7. Estes *SpamBands* utilizam o procoloco SOCKS e mandam muitas mensagens em relação ao número pequeno de endereços IP que possuem, indicando o uso de servidores dedicados para a disseminação de spam.

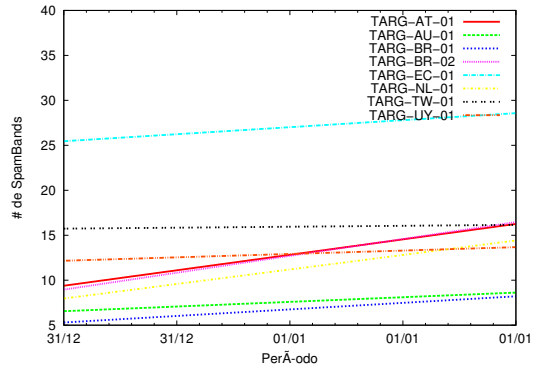
O segundo grupo é formado pelos *SpamBands* 1 e 2, que enviam um número de mensagens muito maior que os demais. Os dois possuem um alto número de endereços IP que estão distribuídos, no *SpamBand* 1, em um AS (3462) do tipo ISP (*Internet Service Provider*) e, no *SpamBand* 2, em 198 ASes. Além disso, grande parte dos endereços IP estão na XBL, o que sugere que estes *SpamBands* podem fazer parte de grandes *botnets* que estão ao redor do mundo e que mandam muito spam.

O terceiro grupo, formado apenas pelo *SpamBand* 5, possui características muito similares aos *SpamBands* 1 e 2, mas o protocolo utilizado é SOCKS. Quatro dos cinco ASes que mais enviaram mensagens neste *SpamBand* são ASes de *hosting*. Isso leva a crer na possibilidade de que algum grupo de disseminação contratou diversos servidores dedicados para enviar suas mensagens.

O *SpamBand* 3 tem características muito similares ao segundo grupo. Entretanto, ele envia um baixo número de mensagens de *spam* e está concentrado em poucos ASes que estão localizados no mesmo *Country Code*. Isso sugere que este *SpamBand* faz parte de uma pequena *botnet*. Já o *SpamBand* 4 traz indícios de que um único serviço de *hosting* está enviando campanhas de spam pela rede.



(a) Número de *SpamBands* por *honeypots* durante o período



(b) Regressão linear do comportamento observado na Fig. 3(a)

Figura 3. Distribuição dos *SpamBands* no período.

B. Visão geral dos SpamBands

A técnica aplicada ao longo de 19 dias gerou um total de 2306 *SpamBands*. A Fig. 3(a) mostra a distribuição desses *SpamBands* ao longo dos dias. Como observado na Tabela II, os dois *honeypots* que mais possuem endereços IP são o *BR-01* e o *EC-01*. Entretanto, observamos no gráfico da

Fig. 3 que esses dois *honeypots* estão em extremos diferentes no gráfico ao longo dos dias, onde o *honeypot EC-01* é o que mais possui *SpamBands* e o *honeypot BR-01*, o que possui menos. Isso sugere que o *honeypot EC-01* é atacado por mais grupos de disseminação de spam do que o *honeypot BR-01*. O restante dos *honeypots* se mantém bem relacionados, mostrando que eles são atacados por um número parecido de grupos de disseminação de spam. A Fig. 3(b) mostra uma regressão linear do número de *SpamBands* por dia por cada *honeypot*. A linha de tendência revela retas com inclinação suave reforçando que a variação observada na Fig. 3(a) tem uma regularidade e representa algum tipo de ofuscação utilizada pelo *spammer*.

Relação entre protocolos

TABELA IV
RELAÇÕES DOS PROTOCOLOS ENTRE *SpamBands*.

	***SpamBands* (%)**
Somente HTTP	12 (0,52)
Somente SMTP	925 (40,10)
Somente SOCKS	891 (38,62)
Somente HTTP e SMTP	1 (0,05)
Somente HTTP e SOCKS	383 (16,60)
Somente SMTP e SOCKS	42 (1,82)
HTTP e SMTP e SOCKS	53 (2,29)

A Tabela IV mostra a distribuição dos protocolos nos *SpamBands*. Através dessa tabela é possível notar uma relação interessante entre HTTP e SOCKS: entre todos os *SpamBands* que utilizam HTTP, 97,10% também utilizam SOCKS. Como ambos protocolos são utilizados para atacar o *honeypot* como *proxy*, isso leva a um forte indício de que o uso desses dois protocolos esteja relacionado com algum tipo de ofuscação, a qual não abordamos mais profundamente neste trabalho.

É possível verificar que muito poucos *SpamBands* utilizam SMTP em conjunto com outro protocolo. Entretanto, existem *SpamBands* que utilizam os protocolos HTTP/SOCKS e SMTP ao mesmo tempo, levando a crer na existência de um grupo de disseminação de spams que utiliza dois ou mais tipos de redes distintas para enviar suas mensagens. Uma possibilidade é o uso tanto de redes *botnets* quanto servidores dedicados para o envio de campanhas de spam. O primeiro tipo de rede tende a utilizar o protocolo SMTP pois o spammer está interessado em apenas repassar suas mensagens, visto que o mesmo já está oculto na rede. No entanto, o uso dos protocolos HTTP e SOCKS, no segundo tipo de rede, indica que o grupo de disseminação de spam utiliza servidores dedicados para o envio de suas mensagens.

Relações entre número de endereços IP, mensagens, CCs e ASes

O gráfico da Fig. 4(a) mostra que apenas 10% dos *SpamBands* com protocolos SOCKS e HTTP têm mais de 100 endereços IP, o que sugere o uso de servidores para o envio. Entretanto, cerca de 37,5% do total de *SpamBands* que possuem o protocolo SMTP têm mais de 100 endereços IP, o que não surpreende, pois redes *botnets*, em geral, são constituídas por um número maior de endereços IP no envio se comparado com HTTP e SOCKS, além de ter como característica o uso do protocolo SMTP. Entretanto, observando a Fig. 4(b) verificamos uma inversão: *SpamBands* HTTP e SOCKS tendem a enviar mais mensagens do que *SpamBands* SMTP. Isso sugere que *SpamBands* SMTP, apesar de serem formados por um grande número de endereços IP, enviam poucas mensagens.

Os gráficos das figuras 4(c) e 4(d) são bastante semelhantes. Aplicando a correlação de Pearson entre o número de *Country codes* e *ASes*, obtemos um coeficiente de 0.95, o que indica que um mesmo *SpamBand* tende a ter comportamento semelhante nos dois gráficos. Dessa forma, a análise para o gráfico 4(c) espelha-se no gráfico 4(d).

O gráfico da Fig. 4(c) sugere que os *SpamBands* que mais estão espalhados pelos países são SMTP, o que mostra uma característica típica de *botnets*. Todavia, cerca de 85% dos *SpamBands* que utilizam o protocolo SMTP contém endereços IP vindos de menos de 10 CCs, o que indica pequenas *botnets*, similar ao *SpamBand* 3 da Tabela III. Por outro lado, todos os *SpamBands* que possuem HTTP e cerca de 90% que possuem SOCKS têm endereços IP de, no máximo, 5 *Country Codes*, indicando grupos de disseminação que utilizam servidores para o envio de suas mensagens. Entretanto, alguns *SpamBands* que usam o protocolo SOCKS (cerca de 10%) chegam a ter mais de 5 *Country Codes*, indicando um comportamento similar ao *SpamBand* 5 da Tabela III.

Interseção de *SpamBands* entre *honeypots*

Como visto na Tabela II, existe uma recorrência de máquinas entre os *honeypots*. Isso leva a crer que um *SpamBand* pode também participar de outros *honeypots*. Para ilustrar essa reincidência, utilizamos os *SpamBands* do *honeypot BR-01* como referência de comparação com *SpamBands* de outros *honeypots*. A Fig. 5 apresenta essa visão.

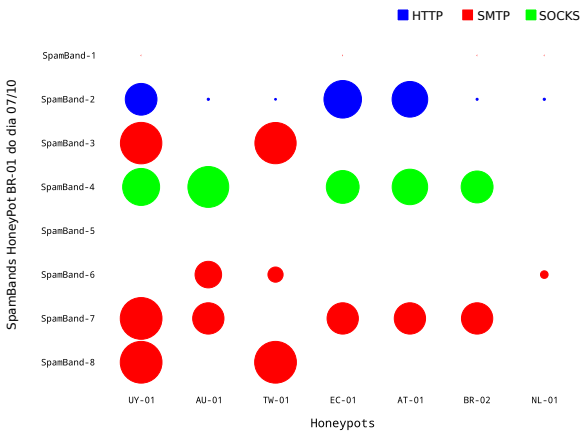
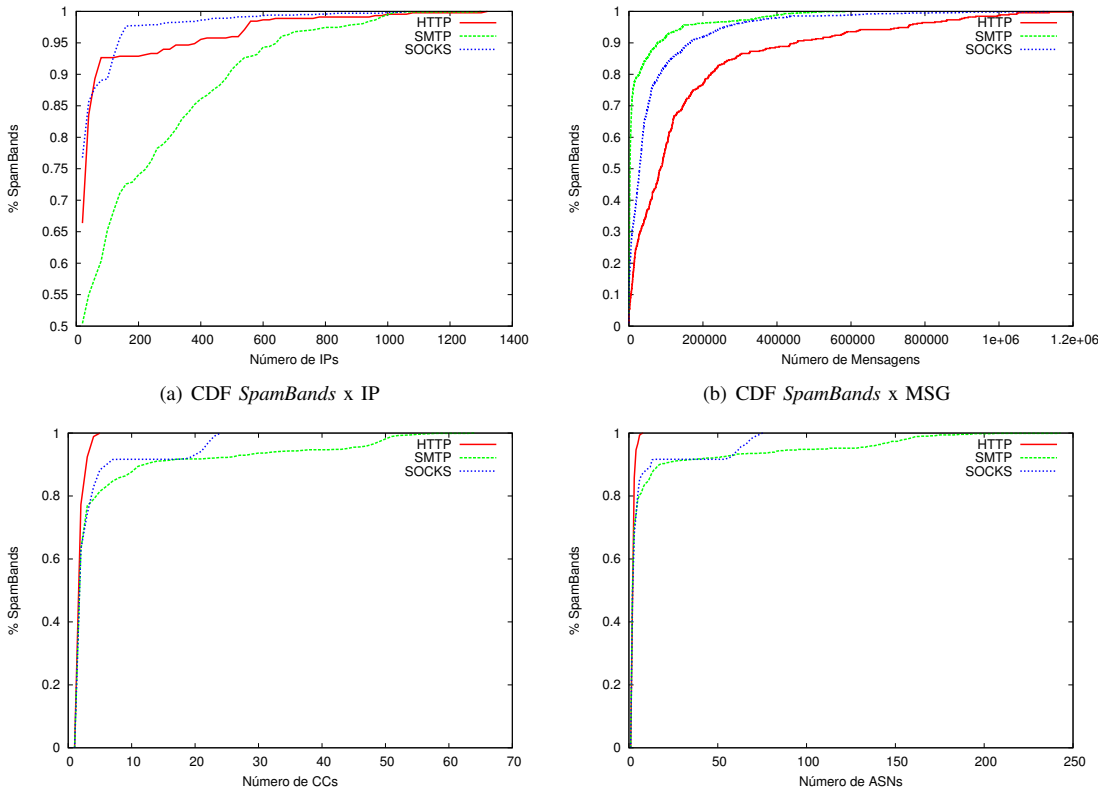


Figura 5. Interseção dos *SpamBands* do *honeypot BR-01* com *SpamBands* de outros *honeypots*.

Analisando a figura, é possível verificar que se máquinas de algum *SpamBand* aparecem em outros *honeypots*, elas



(a) CDF *SpamBands* x IP

(b) CDF *SpamBands* x MSG

(c) CDF *SpamBands* x CC

(d) CDF *SpamBands* x AS

Figura 4. *SpamBands* em relação ao número de endereços IP, mensagens, CCs e ASes.

tendem a estar no mesmo *SpamBand*. O *SpamBand* 5 do *honeypot BR-01*, não possui máquinas em outros *honeypots*, o que indica que esse grupo de máquinas têm visão apenas do *honeypot* usado como referência. Esse fato se assemelha com o *SpamBand* 1, que é o maior em número de máquinas. Entretanto, esse *SpamBand* possui uma única máquina nos *honeypots UY-01, EC-01, BR-02 e NL-01*. Isto leva a crer que o *SpamBand* possui conhecimento dos *honeypots* citados mas, por algum motivo desconhecido, está utilizando apenas o *honeypot BR-01*.

Averiguando os *SpamBands* 2, 3, 4, 6, 7 e 8, vê-se que estes grupos conseguem alcançar outros *honeypots*. Além disso, eles não têm recorrência nos mesmos *honeypots*, o que reforça a hipótese de estes grupos serem independentes. Outro fato importante é que esses *SpamBands* também não utilizam todas as máquinas em todos os *honeypots*. Essa evidência leva a crer que existe algum tipo de distribuição de atividades desses *SpamBands* na rede.

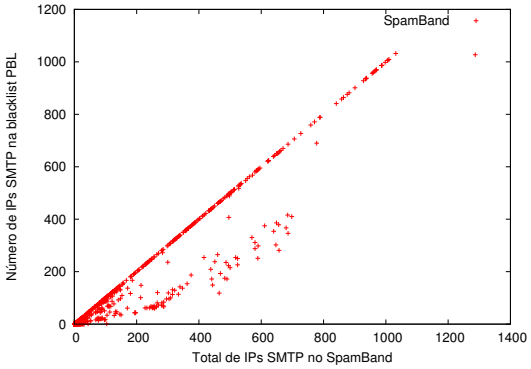
C. *Relação entre SpamBands e* blacklists

A Tabela V fornece a relação entre o número de IPs dos *SpamBands* que estão em *blacklists* e o número de IPs que o *SpamBand* possui em cada protocolo. Observamos uma correlação muito forte no número de IPs na *PBL (Policy Block List)* e o número de IPs do protocolo SMTP nos *SpamBands*. Conforme observado na seção IV-B, 90% dos *SpamBands* possuem somente o protocolo SMTP, o que leva a um forte indício desses *SpamBands* serem partes de *botnets*. Pela Fig. 6(a) observamos que a *PBL* captura grande parte desses IPs que estão nos *SpamBands* e que possivelmente fazem parte de *botnets*. Por outro lado, os protocolos HTTP e SOCKS possuem uma correlação fraca, o que era esperado visto que *SpamBands* desse tipo tendem a enviar suas mensagens de serviços de *hosting*.

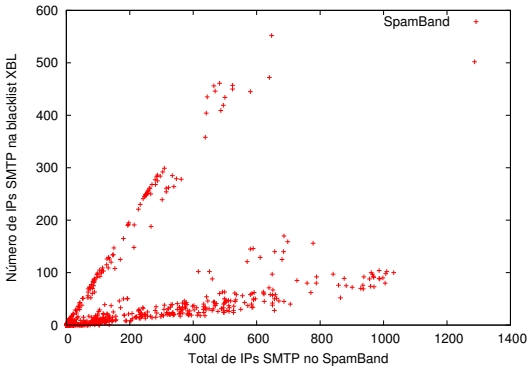
TABELA V
COEFICIENTE DE DETERMINAÇÃO ENTRE PROTOCOLOS DOS *SpamBands* E *Blacklists*.

	PBL	**XBL**
HTTP	0.38	0.11
SMTP	0.86	0.55
SOCKS	0.35	0.08

Em relação a *XBL (Spamhaus Exploits List)*, observamos uma correlação moderada, o que não é esperado visto que diversas máquinas que estão em *botnets* estão infectadas por algum tipo de *malware*. Analisando o gráfico da Fig. 6(b) observamos o porquê da relação ter sido moderada: existem dois eixos de tendência. O primeiro é uma relação linear entre o número de IPs na XBL do *SpamBand* e o número de IPs SMTP, que seria esperado: os endereços IP de todos os participantes de uma *botnet* tendem a acabar sendo identificados por *blacklists*. Entretando, o segundo eixo possui uma relação 1:10, o que sugere algum comportamento especial por parte daqueles *SpamBands*. Eles não só conseguem mascarar bem as atividades de suas máquinas na rede do ponto de vista



(a) Distribuição de IPs SMTP por IPs SMTP na PBL nos *SpamBands*



(b) Distribuição de IPs SMTP por IPs SMTP na XBL nos *SpamBands*

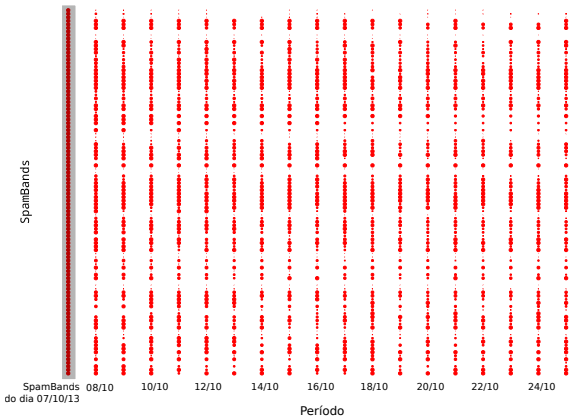
Figura 6. Relação entre o protocolo SMTP e as *blacklists PBL* e *XBL* nos *SpamBands*.

das *blacklists*, mas mantêm uma taxa comum de identificação em tais listas, o que sugere um comportamento planejado. Determinar a razão para tal comportamento, entretanto, exige novas análises e coletas, sendo considerada como trabalho futuro.

D. Relação Temporal

Nesta seção, procuramos entender o comportamento dos *SpamBands* do primeiro dia do período avaliado (07/10/2013) em outros dias. O método utilizado para verificar a continuidade do *SpamBand* é recuperar o *SpamBand* do mesmo *honeypot*, no dia seguinte, que mais possui IPs em comum com o *SpamBand* do dia de referência. Observe que esta técnica permite que novos IPs apareçam no *SpamBand* ao longo dos dias e que iremos discutir mais adiante. A Fig. 7 mostra que existe uma tendência dos *SpamBands* permanecerem ao longo do tempo. O tamanho dos pontos do gráfico indicam quantos IPs permaneceram em relação ao dia de referência.

Pode-se notar pelo gráfico da Fig. 7 que os *SpamBands* mudam constantemente seu tamanho ao longo dos dias. Para uma visão geral do comportamento temporal dos *SpamBands* por protocolo, procuramos observar dois quesitos: a variação do tamanho e a estabilidade dos IPs que participam do *SpamBand* no período avaliado. O primeiro quesito é calculado através do coeficiente de variação e o segundo, dividindo a média



(a) Comportamento dos *SpamBands* do dia 07/10/2013 ao longo dos dias

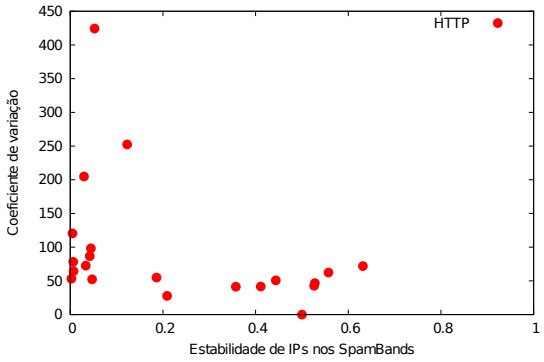
Figura 7. Comportamento geral dos *SpamBands* ao longo dos dias.

de IPs pelo número total de IPs distintos que apareceram no período. A Fig. 8 mostra uma relação global entre protocolos, estabilidade e variação dos *SpamBands*. Pelas figuras 8(a) e 8(c), observamos que os *SpamBands* HTTP e SOCKS tendem a manter seu tamanho e possuir maior estabilidade. Isso reforça, mais uma vez, que os *SpamBands* baseados nesses protocolos utilizam serviços de *hosting* para enviar suas mensagens. Por outro lado, vemos um comportamento diferenciado do protocolo SMTP na Fig. 8(b), que indica que esses *SpamBands* são bem menos estáveis que os dos protocolos HTTP e SOCKS e possuem maiores variações no tamanho, o que indica uma dinamicidade nesses *SpamBands*.

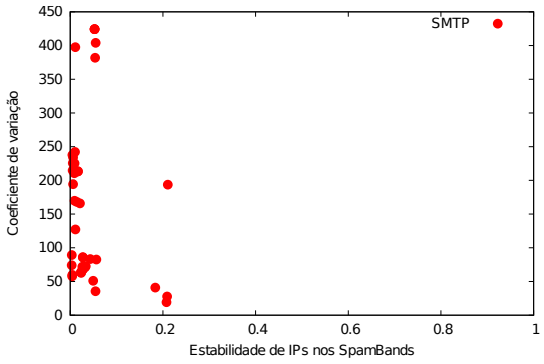
Exemplo de relação temporal entre campanhas e IPs nos *SpamBands*

Para ilustrar o comportamento das campanhas em relação a mudança dos IPs nos *SpamBands*, realizamos o mesmo método aplicado anteriormente para identificar *SpamBands* semelhantes, entre dias, através de IPs. Entretanto, utilizamos campanhas ao invés de IPs. Nos dois gráficos da Fig. 9, mostramos uma relação entre os IPs e campanhas nos *SpamBands*. A Fig. 9 mostra o experimento realizado para o *honeypot BR-01*. Como podemos observar na figura, existe uma relação entre o comportamento dos grupos de IPs e campanhas. O *SpamBand* 6 do *honeypot BR-01* no dia 07/10/2013 desaparece completamente no dia 12/10/2013, como mostra a Fig. 9(a). Todos os 47 IPs deste *SpamBand* são SMTP e fazem parte de apenas um *country code* (CN) e um AS (4134) do tipo DSL (4134), o que indica um *SpamBand* que faz parte de uma pequena *botnet*. Como os IPs do tipo DSL tendem a ser dinâmicos, é possível essas máquinas finais tenham mudado seu endereço IP durante os dias. Entretanto, a possibilidade delas terem saído da *botnet* é maior visto que as campanhas que elas suportavam também desapareceram.

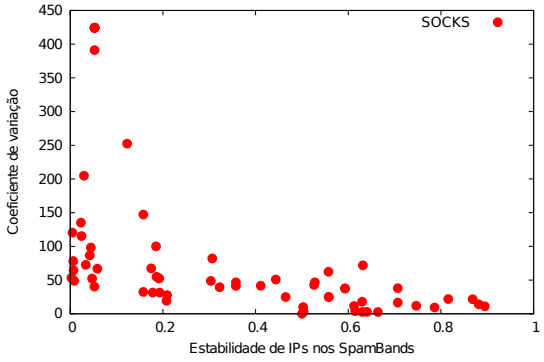
O *SpamBand* 5 retrata um grupo puramente SOCKS, onde IPs estão distribuídos em 23 *country codes* e 72 ASes. Esse grupo é semelhante ao *SpamBand* 5 do estudo de caso da Seção IV-A, que possivelmente contratou diversos serviços



(a) Variação x Estabilidade HTTP



(b) Variação x Estabilidade SMTP

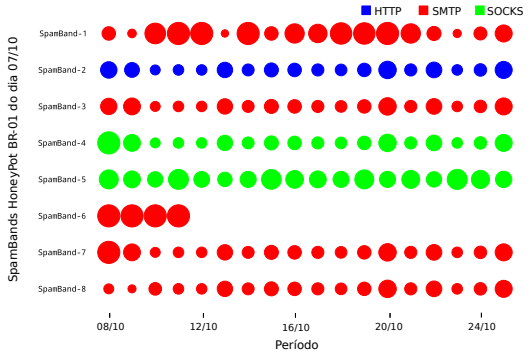


(c) Variação x Estabilidade SOCKS

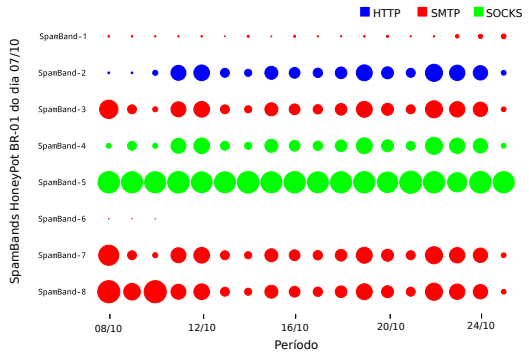
Figura 8. Estabilidade e desvio padrão relativo de IPs nos *SpamBands* , ao longo do dia, por protocolo.

de *hosting* para o envio das campanhas. Observe que o grupo de máquinas permanece similar ao longo do tempo e existe uma periodicidade no grupo de campanhas, mostrando que este grupo de máquinas enviam um mesmo grupo de campanhas alternadas ao longo dos dias avaliados.

Um outro comportamento interessante que podemos observar é sobre os *SpamBands* 2,3,4,7 e 8. No dia 11/10, estes *SpamBands* se unem, formando um único *SpamBand* e, por isso, o comportamento temporal desses 5 *SpamBands* nos dois gráficos da Fig. 9 são iguais. Este fato indica que estes *SpamBands* podem estar oferecendo serviços que são contratados por grupos de disseminação e que, em algum momento, um grupo contratou estes serviços para enviarem as mesmas campanhas.



(a) Comportamento dos IPs dos *SpamBands* do *honeypot BR-01* do dia 07/10 ao longo do período analisado.



(b) Comportamento das campanhas nos *SpamBands* do *honeypot BR-01* do dia 07/10 ao longo do período analisado.

Figura 9. Exemplo de relação entre campanha e IPs nos *SpamBands*.

V. Trabalhos Relacionados

Alguns autores já focaram no comportamento dos *spammers* de formas que tiveram impacto sobre este trabalho.

Guerra et al. apresentam uma técnica que utiliza uma estrutura de mineração de dados denominada FPTree para agrupar mensagens de spam [8]. As mensagens assim agrupadas definem o conceito de campanha de spam, como usadas neste trabalho: uma campanha é um conjunto de mensagens que foram enviadas com um mesmo objetivo mas que foram diferenciadas por algum tipo de ofuscação, com a finalidade de não serem captadas por filtros spam.

Ramachandran et al. mostram que o *spammer* alterna as máquinas usadas para envio, de modo ocultar sua origem e contornar diversos filtros de spam na rede [9]. Esse trabalho sugere que as mensagens de uma mesma campanha podem ser enviadas por diferentes máquinas, o que motiva nosso trabalho para encontrar uma forma de agrupar essas máquinas.

Moreira Moura et al. introduz o conceito de *Bad Neighborhoods*, que são vizinhanças de rede com alta probabilidade de um IP enviar spam [10]. *Fonseca et al.* estende esse conceito e estabelece uma relação direta de vizinhança com Sistemas Autônomos (AS), por esses terem fronteiras bem definidas. Nosso trabalho apresenta uma visão complementar a esses conceitos: ao invés de focarmos diretamente nos locais de

origem do *spam*, estamos procurando entender como diferentes origens (máquinas em diferentes pontos da rede) se relacionam para atender às necessidades do *spammer*, o orquestrador por trás de todo o processo.

Por fim, Zhuang *et al.* associa características de spam a *botnets*, que são um meio de envio de mensagens de spam [11]. Contudo, existe a possibilidade de que vários *spammers* utilizem a mesma *botnet* ou combinações delas, visto que essas redes muitas vezes são alugadas para terceiros [12]. Dessa forma, sem uma identificação das campanhas de spam sendo enviadas, grupos de máquinas utilizadas por diferentes spammers podem ser vistas como uma só entidade, não levando a um bom agrupamento de máquinas.

VI. Conclusão e Trabalhos Futuros

Neste trabalho, buscamos entender melhor o comportamento dos *spammers* correlacionando as máquinas utilizadas para envio através das campanhas de spam enviadas. Para realizar essa análise, propusemos o conceito de *SpamBands*, grupos de máquinas que participam das mesmas campanhas e sugerem a existência de um único orquestrador por trás de seu comportamento e desenvolvemos uma metodologia baseada em grafos para identificar esse grupos. Inicialmente, conectamos todas as máquinas que enviam as mesmas campanhas. Os grupos revelados por esta metodologia passam por um processo de refinamento, de forma a separar subgrafos densos, que revelam os *SpamBands*.

Descobrimos que a grande maioria dos *SpamBands* tendem a utilizar apenas o protocolo SMTP ou os protocolos HTTP/SOCKS, o que faz uma distinção entre grupos que utilizam servidores dedicados e redes *botnets* para o envio de spam. Além disso, mostramos que esse conceito permite revelar grupos que não são inteiramente detectados pela *blacklist XBL*, podendo ajudar a inferir outras máquinas que deveriam pertencer à *blacklist*. Ademais, encontramos *SpamBands* que utilizam os dois tipos de protocolos, levando a crer na existência de grupos de disseminação de spam que utilizam tanto servidores dedicados quanto redes *botnets* para enviar mensagens.

Por fim, realizamos ainda um estudo sobre os *SpamBands* ao longo dos dias avaliados, revelando que eles se repetem ao longo do tempo. Nesta avaliação, descobrimos que *SpamBands* que utilizam os protocolos HTTP/SOCKS tendem a ser mais estáveis em relação ao número de IPs, o que não acontece com *SpamBands* que utilizam o protocolo SMTP. Como trabalho futuro, pretendemos analisar mais profundamente o comportamento dos *SpamBands* ao longo dos dias, buscando a existência de uma interação entre eles, de forma a entender o comportamento temporal.

Agradecimentos

Este trabalho foi parcialmente financiado por NIC.BR, Fapemig, CAPES, CNPq e InWeb.

Referências

[1] D. Crocker, "Challenges in anti-spam efforts," *The Internet Protocol Journal*, vol. 8, no. 4, 2006. [Online]. Available: "http://www.cisco.com/web/about/ac123/ac147/archived_issues/ipj_8-4/anti-spam_efforts.html"

[2] Royal Pingdom, "The internet 2012 in numbers," Artigo na Web, Visitado em 2014. [Online]. Available: http://royal.pingdom.com/2013/01/16/internet-2012-in-numbers/

[3] J. C. Sipior, B. T. Ward, and P. G. Bonner, "Should spam be on the menu?" *Commun. ACM*, vol. 47, no. 6, pp. 59–63, Jun. 2004. [Online]. Available: http://doi.acm.org/10.1145/990680.990681

[4] G. V. Cormack, "Email spam filtering: A systematic review," *Found. Trends Inf. Retr.*, vol. 1, no. 4, pp. 335–455, Apr. 2008. [Online]. Available: http://dx.doi.org/10.1561/1500000006

[5] P. H. B. Las-Casas, D. Guedes, W. M. Jr., C. Hoepers, K. Steding-Jessen, M. H. P. Chaves, O. Fonseca, E. Fazzion, and R. E. A. Moreira, "Análise do tráfego de spam coletado ao redor do mundo," in *Anais do Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos (SBRC)*. SBC, 2013.

[6] P. H. C. Guerra, D. Guedes, W. M. Jr., C. Hoepers, and K. Steding-Jessen, "Caracterização de estratégias de disseminação de spams," in *Anais do Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos (SBRC)*. SBC, 2008.

[7] H. Almeida, D. Guedes, W. Meira, and M. J. Zaki, "Is there a best quality metric for graph clusters?" in *Proceedings of the 2011 European Conference on Machine Learning and Knowledge Discovery in Databases - Volume Part I*, Athens, Greece, 2011, pp. 44–59.

[8] P. H. C. Guerra, D. E. V. Pires, D. Guedes, J. Wagner Meira, C. Hoepers, and K. Steding-Jessen, "A campaign-based characterization of spamming strategies," in *Proceedings of the 5th Conference on e-mail and anti-spam (CEAS)*, Mountain View, CA, 2008.

[9] A. Ramachandran and N. Feamster, "Understanding the network-level behavior of spammers," *SIGCOMM Comput. Commun. Rev.*, vol. 36, no. 4, pp. 291–302, Aug. 2006. [Online]. Available: http://doi.acm.org/10.1145/1151659.1159947

[10] G. C. Moreira Moura, R. Sadre, and A. Pras, "Internet bad neighborhoods: the spam case," in *7th International Conference on Network and Services Management (CNSM 2011), Paris, France*, O. Festor and E. Lupu, Eds. USA: IEEE Communications Society, October 2011, pp. 1–8.

[11] L. Zhuang, J. Dunagan, D. R. Simon, H. J. Wang, I. Osipkov, and J. D. Tygar, "Characterizing botnets from email spam records." in *LEET*, F. Monrose, Ed. USENIX Association, 2008.

[12] D. Raywood, "The botnet market and what you get for your money," *SC Magazine UK*, 2010.

Elverton Fazzion é aluno de mestrado do Departamento de Ciência da Computação da Universidade Federal de Minas Gerais. Possui graduação em Ciência da computação pela Universidade Federal de Minas Gerais (2014). Seus interesses são na área de redes de computadores, mineração de dados e algoritmos.

Pedro Las-Casas é aluno de doutorado do Departamento de Ciência da Computação da Universidade Federal de Minas Gerais. Possui graduação em Ciência da Computação pela PUC Minas (2010) e mestrado pela UFMG (2013). Seus interesses são em redes de computadores, processamento massivo de dados e mineração de dados.

Osvaldo Fonseca é aluno de mestrado do Departamento de Ciência da Computação da Universidade Federal de Minas Gerais. Possui graduação em Ciência da computação pela Universidade Federal de Minas Gerais (2013). Seus interesses são na área de redes de computadores e mineração de dados.

Dorgival Guedes possui graduação e mestrado em Ciências da Computação pela Universidade Federal de Minas Gerais e doutorado pela University of Arizona, Tucson (1999). É professor associado do departamento de Ciência da Computação da UFMG, Brasil. Atuou como professor visitante no International Computer Science Institute (ICSI) e na University of California, Berkeley, em 2011. Suas áreas de pesquisa incluem Redes de Computadores, Sistemas Distribuídos e Sistemas Operacionais, especialmente quando elas são relacionadas com escalabilidade de aplicações distribuídas, incluindo áreas como Computação em Núvem, Big-Data, e Redes Definidas por Software.

Wagner Meira Jr. Bacharel e mestre em Ciência da Computação pela Universidade Federal de Minas Gerais (1990 e 1993) e doutor em Ciência da Computação pela University of Rochester (1997). Já publicou mais de 200 artigos em veículos de comunicação de grande importância e é co-autor do livro *Data Mining and Analysis - Fundamental Concepts and Algorithms* publicado pela Cambridge University Press em 2014. Atualmente é professor titular da Universidade Federal de Minas Gerais. Suas áreas de interesse são sistemas paralelos e distribuídos e mineração de dados, assim como a sua aplicação em redes sociais, comércio eletrônico, recuperação de informação e bioinformática, entre outros.

Cristine Hoepers possui graduação em Ciências da Computação pela Universidade Federal de Santa Catarina (1996) e doutorado em Computação Aplicada pelo Instituto Nacional de Pesquisas Espaciais (2008). É Gerente Geral do CERT.br/NIC.br, onde está desde 1999, e atua no desenvolvimento de boas práticas de segurança, na conscientização de usuários e na coordenação do honeyTARG Project, Capítulo do Honeynet Project Mundial. Tem experiência nas áreas de Redes de Computadores, Segurança, Gestão de Incidentes e uso de Honeypots para Análise de Tendências e Detecção de Ataques.

Klaus Steding-Jessen possui graduação em Engenharia da Computação pela UNICAMP (1996) e doutorado em Computação Aplicada pelo Instituto Nacional de Pesquisas Espaciais (2008). É Gerente Técnico do CERT.br/NIC.br, onde está desde 1999, atuando nas áreas de infraestrutura, treinamento e análise de tendências, esta última como parte do honeyTARG Project, capítulo do Honeynet Project Mundial. Tem experiência nas áreas de Redes de Computadores, Segurança, Tratamento de Incidentes e uso de Honeypots para Análise de Tendências e Detecção de Ataques.

Marcelo Chaves é bacharel em Ciência da Computação pela Universidade Federal de Ouro Preto (1999) e mestre em Computação Aplicada pelo Instituto Nacional de Pesquisas Espaciais (2002). É analista de projetos de segurança senior do CERT.br/NIC.br, onde está desde 2002. Atua na área de infraestrutura, pesquisa e desenvolvimento do CERT.br, além de ser membro e desenvolvedor do honeyTARG Project, capítulo do Honeynet Project Mundial. Suas especialidades incluem tecnologias envolvendo honeypots, análise de incidentes (incluindo fraudes via Internet), metodologias antispam, arquiteturas de segurança, monitoramento de redes e análise de logs.