Identificação e Caracterização de Spammers a partir de Listas de Destinatários

Pedro H. Calais Guerra[1], Marco Túlio Ribeiro[1], Dorgival Olavo Guedes[1], Wagner Meira Jr.[1]
Cristine Hoepers[2], Klaus Steding-Jessen[2], Marcelo H. P. C. Chaves[2]

[1]Departamento de Ciência da Computação – Universidade Federal de Minas Gerais
Belo Horizonte, MG.

[2]CERT.br - Centro de Estudos, Resposta e Tratamento de Incidentes de Segurança no Brasil
NIC.br - Núcleo de Informação e Coordenação do Ponto br, São Paulo, SP

{pcalais,marcotcr,dorgival,meira}@dcc.ufmg.br

{cristine,jessen,mhp}@cert.br

Abstract. *In this work, we analyze spamming dissemination patterns that can be inferred from spam's recipient lists. From the set of recipients associated to each source IP address, we determined the sets of IP numbers that systematically abuse the same targets. We then caracterized how those sets abuse their recipient lists and unveil properties that may be employed as spam mitigation criteria. Our study unveils that up to 100% of the recipients targeted by a given IP address are also targeted by other IP addresses geographically close within the network and that it is possible to determine spammers considering only spam source and targets.*

Resumo. *Neste trabalho, analisamos padrões de disseminação de* spams *que podem ser inferidos a partir das listas de destinatários abusadas por* spammers. *A partir da determinação de conjuntos de endereços IP que abusam consistentemente os mesmos destinatários, caracterizamos propriedades da disseminação de listas de destinatários desses conjuntos que podem ser utilizadas como critérios para identificação e detecção de* spams. *Nosso estudo revela que até 100% dos destinatários abusados por um endereço IP são também alvo de endereços IPs geograficamente próximos na rede, e que é possível isolar grupos de* spammers *considerando apenas a origem e o destino das mensagens de* spam.

1. Introdução

O *spam* é um dos maiores problemas de abuso da infraestrutura da Internet [Hayes 2003, Cranor and LaMacchia 1998]. Relatórios recentes apontam que entre 82% e 92% das mensagens recebidas por seus servidores são *spams* [MAAWG 2009] e o prejuízo que essa prática acarreta às empresas e à sociedade em geral é avaliado em vários bilhões de dólares [Sipior et al. 2004].

O combate ao *spam* se caracteriza por uma evolução constante das técnicas de detecção das mensagens indesejadas, motivada pelo aumento da sofisticação das tecnologias adotadas pelos próprios *spammers* [Goodman et al. 2007]. Essa evolução acontece tanto no modo como disseminam suas mensagens pela rede, buscando maximizar o volume de mensagens que enviam enquanto mantêm sua identidade oculta, quanto na forma como constróem o conteúdo das mensagens [Pu and Webb 2006].

No que se refere à forma de atuação na rede, a tendência mais clara que pode ser percebida é a de que os originadores de *spam* primam por dispersar ao máximo seus abusos [Kokkodis and Faloutsos 2009, Calais et al. 2009], dificultando enormemente a identificação individual de cada emissor de *spam*, seja ele um *bot* ou uma máquina que se comporta como *proxy* ou *mail relay* aberto. O fato da maior parte das máquinas de usuários comprometidas para enviar *spam* ser identificada na rede por endereços IP dinâmicos [Xie et al. 2007] torna o problema ainda mais desafiador. Um fenômeno similar ocorre na geração do conteúdo dos *spams*: evita-se ao máximo gerar mensagens de conteúdo idêntico, o que facilitaria a identificação dos abusos. Inserir aleatoriedade nas mensagens é uma prática comum entre *spammers*.

Por esses motivos, buscar critérios que consigam identificar similaridades entre as diferentes origens do *spam* é um caminho relevante para melhorar as técnicas de bloqueio de *spams* ou conseguir, ao menos, manter a efetividade do combate às mensagens indesejadas [Li and Hsieh 2006]. Em relação à ofuscação das mensagens, existem técnicas que tentam desfazer esse processo a partir da identificação de *campanhas* [Calais et al. 2008, Yeh and Lin 2006], que são grupos de mensagens que, embora sejam diferentes, possuem um alto grau de similaridade e têm a mesma finalidade. A identificação de campanhas já se mostrou uma técnica eficaz para identificação de *spams* e caracterização da atuação de *spammers* [Calais et al. 2008]. Para fins de caracterização, a principal aplicação da identificação de campanhas é fornecer um critério de sumarização dos *spams* e revelar estratégias de geração do conteúdo dessas mensagens.

No entanto, campanhas de *spam*, em geral, apresentam curta duração e por isso são insuficientes para caracterização de padrões de longo prazo de *spammers*. Neste trabalho exploramos uma dimensão, complementar e ortogonal às campanhas, que permite analisar novas características do tráfego de *spam*: os destinatários das mensagens. A principal motivação para analisar o conjunto de destinatários alvo dos *spams* é que ele é uma característica marcante de um *spammer*: por mais que ele altere sua origem na rede e o conteúdo de suas campanhas de *spam*, o destinatário das mensagens não pode ser ofuscado, sob pena dos servidores SMTP responsáveis pela entrega não serem capazes de identificá-lo. Os destinatários, portanto, tendem a ser um importante *invariante* que identifica e representa um *spammer* individual ou conjunto de *spammers*. O objetivo deste trabalho é explorar essa dimensão e mostrar seu potencial em identificar *spammers* em um grão mais grosso que endereços IP e, possivelmente, colaborar para o estabelecimento de novos critérios de detecção e mitigação do *spam*. Em particular, a análise das listas de destinatários permite observar aspectos de *intensidade* e *dispersão* dos abusos que a análise de campanhas isoladamente não permite.

A metodologia do trabalho consiste em avaliar as relações entre os endereços IP de origem de *spam* e os destinatários visados por eles, bem como as interseções entre os destinatários utilizados por diferentes origens. A nossa proposta para identificar e caracterizar *spammers* a partir de listas de destinatários consiste em três etapas bem definidas. A primeira etapa refere-se à coleta dos dados, realizada a partir de *honeypots* de baixa interatividade. Em seguida, extraímos o endereço IP de origem de cada mensagem, assim como seus destinatários, e agrupamos os endereços IP de acordo com a similaridade de suas listas de destinatários. Os grupos de endereços IP determinados representam diferentes *spammers* (ou grupo de *spammers* que atuam sob a mesma lista). A terceira parte da

metodologia consiste em analisar cada um dos grupos em termos de suas campanhas de *spam*, diversidades de origens e tamanho de listas de destinatários, entre outros critérios.

O trabalho está organizado como se segue. A seção 2 detalha a estratégia de coleta. Na seção 3, é apresentada a estrátegia para encontrar *spammers* a partir de suas listas de destinatários, cujas propriedades são estudadas na seção 4. Finalmente, os trabalhos relacionados e as conclusões são apresentadas.

2. Coleta dos Dados

A captura das mensagens de *spam* consideradas neste trabalho foi realizada utilizando *honeypots* de baixa interatividade. Os *honeypots* foram configurados de modo a simular computadores com *proxies* e *mail relays* abertos, tradicionalmente abusados para o envio de *spam* e para a realização de outras atividades maliciosas [Krawetz 2004].

Duas máquinas foram instaladas em redes brasileiras e outras quatro nos seguintes países: Áustria, Holanda, Estados Unidos e Uruguai. O objetivo de implantar *honeypots* em mais de um país é obter uma visão global do problema do *spam*, livre de distorções dos dados que poderiam ocorrer ao coletar *spams* apenas em um determinado local.

A captura de mensagens utilizou o *software* `honeyd` [Provos and Holz 2007] em conjunto com subsistemas de emulação de SMTP e *proxies* HTTP e SOCKS desenvolvidos para esse fim. Qualquer máquina que se conectasse à porta 25 de um dos *honeypots* teria a impressão de estar interagindo com um servidor SMTP configurado como *relay aberto*, pronto a repassar as mensagens. Já máquinas que se conectassem a portas tradicionais de *proxies* abertos seriam levadas a acreditar que suas conexões a servidores SMTP externos seriam bem-sucedidas. Todas as transações efetuadas pelos subsistemas do `honeyd` foram armazenadas em *logs* com informações como data e hora, IP de origem da atividade e protocolo que foi abusado no *honeypot*. No caso do abuso ser a um *proxy*, foi registrado também o endereço IP e porta de destino pretendidos como alvo da conexão. Todas as mensagens SMTP observadas, seja por terem sido entregues ao *relay* ou por terem passado pelos *proxies*, foram armazenadas com endereços de destino e conteúdo.

Outras informações eram obtidas através de processamento desses dados armazenados através de consultas ao DNS e outras bases de dados apropriadas. De interesse particular para o objetivo deste trabalho, a partir dos endereços IP de origem das conexões foram também obtidos o identificador do país de origem da conexão (*country code*, CC, obtido através das informações de alocação de enderecos IP mantidas nos arquivos de estatísticas dos 5 Regional Internet Registries (RIR): AfriNIC, APNIC, ARIN, LACNIC e RIPE NCC) e o prefixo de rede onde o endereço estava contido. O endereço de correio eletrônico de cada destinatário era processado para extrair o nome do domínio associado.

Os *honeypots* proveem uma visão de "dentro" da rede, coletando *spams* entre alguma origem anterior do processo de entrega da mensagem e o destino final. Assim, conseguimos observar vários domínios de correio eletrônico sendo visados como destino da mensagens, o que nos garante um ponto de vista global [Pathak et al. 2008] para análise do *spam*. Essa perspectiva fora de qualquer domínio específico é chave para analisar padrões de disseminação de listas de destinatários.

A tabela 2 exibe os números gerais da coleta realizada entre abril e outubro de 2009. Durante esses 5 meses, cerca de 600 milhões de endereços destinatários foram

alvos de 40 milhões de *spams* (uma mensagem pode ser entregue a vários destinatários), dentre os quais 97,4 milhões de destinatários são únicos. Dessa forma, cada destinatário receberia, em média, cerca de 6 mensagens.

Tabela 1. Sumário geral dos Dados Utilizados no Trabalho

período de análise	13/05/2009 a 09/10/2009
número de mensagens	40.026.883
endereços IP de origem distintos	49.957
destinatários das mensagens	597.321.121
destinatários distintos	97.421.912
domínios de correio destino distintos	69.286

Uma vez de posse dos dados de todas as conexões, cada endereço de origem distinto foi identificado e todas as mensagens originadas daquela origem foram processadas. Cada endereço de destinatário encontrado nessas mensagens era inserido em uma lista para aquele endereço IP de origem. Repetições de um mesmo destinatário são contadas para se obter o total de endereços que foram alvos de mensagens, bem como o total de destinatários distintos. Além disso, mensagens de teste comumente enviadas por *spammers* para encontrar na rede máquinas vulneráveis e que podem ser abusadas são identificadas pelos *honeypots*. Essas mensagens, em geral, contém características próprias como o endereço IP que o *spammer* deseja testar no campo *assunto* da mensagem.

3. Agrupamento de Máquinas Disseminadoras de *Spam*

A partir da impressão inicial de que há uma repetição razoável de destinatários, o próximo passo da metodologia consiste em um método que encontre grupos de origens de *spams* que abusem os mesmos destinatários. Neste trabalho, consideramos como unidade disseminadora de *spam* um endereço IP, mas outras granularidades (como sub-rede e Sistema Autônomo) poderiam ser adotadas.

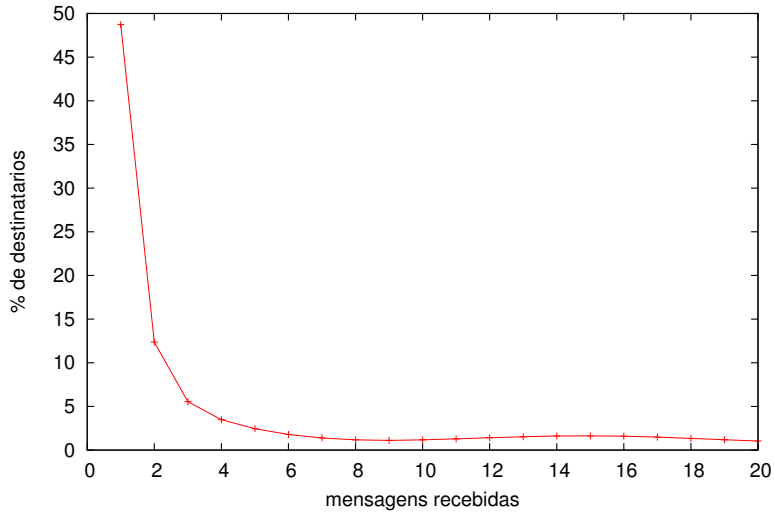

Figura 1. número de mensagens recebidas por cada destinatário

A figura 1 mostra o grau de reincidência dos destinatários no conjunto de dados analisado. Embora uma grande parte dos destinatários (48%) tenha recebido apenas uma

mensagem, a maioria dos endereços de correio recebeu entre duas e vinte mensagens. Essa proporção é similar à apontada em outros trabalhos [Prince et al. 2005], que utilizaram metodologias de coleta diferentes. Existe uma pequena parcela de destinatários, suprimida da figura, que recebeu um volume ainda maior de mensagens (> 100). Esses destinatários correspondem aos *e-mails* dos próprios *spammers*, que enviam mensagens a si próprios para testar o envio de suas mensagens, e foi possível confirmar esses casos porque essas mensagens são especialmente tratadas em nossos *honeypots*, conforme explicado na Seção 2.

Para entender melhor a repetição dos destinatários no conjunto de dados analisado, verificamos a proporção de destinatários associados a *cada* endereço IP. A figura 2 exibe esta proporção para os 300 IPs que mais tentaram enviar mensagens por meio dos *honeypots*, responsáveis por 88% do tráfego observado. Os endereços estão ordenados pela quantidade de destinatários abusados.

É possível observar que os endereços IP que mais enviam mensagens tendem a repetir os mesmos destinatários. No entanto, a maior parte dos endereços IP repetem poucas vezes seus alvos.

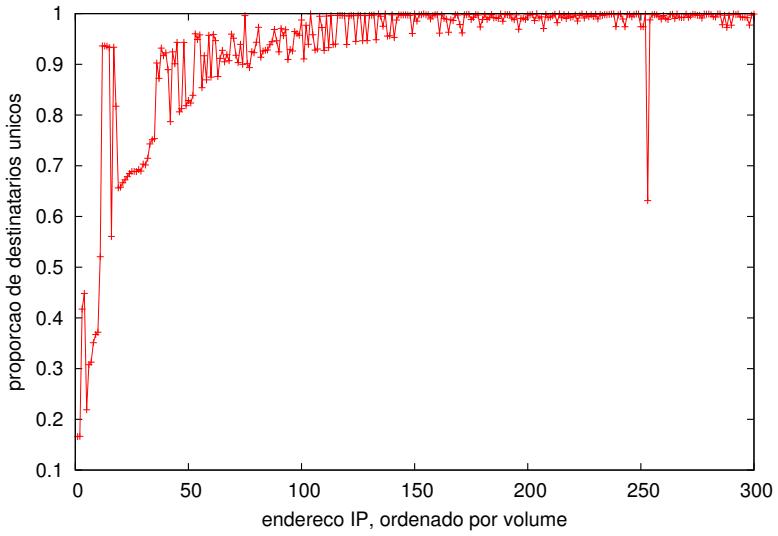

Figura 2. proporção de destinatários únicos

Verificamos, então, a sobreposição de destinatários entre diferentes endereços IP, isto é, a proporção de destinatários que um endereço IP tentou atingir que foi encontrada em outros endereços IP. Para nossa surpresa, esta proporção é bastante alta: em média, 72% dos destinatários alvos de um endereço IP são encontrados em algum outro (ou outros) endereços. A figura 3 exibe um histograma onde a porcentagem de compartilhamento de cada endereço IP com outros endereços é mostrada. Alguns deles, inclusive, chegam a compartilhar mais de 90% de seus endereços com outros endereços. No conjunto de dados analisado, apenas 0,2% dos endereços IP não compartilharam destinatários com nenhuma outra origem, e 9,8% dos pares de endereços IP compartilham pelo menos um destinatário entre si.

Esta sobreposição motivou o projeto de um algoritmo para determinar *listas de destinatários*, isto é, conjuntos de endereços que são abusados pelos mesmos conjuntos

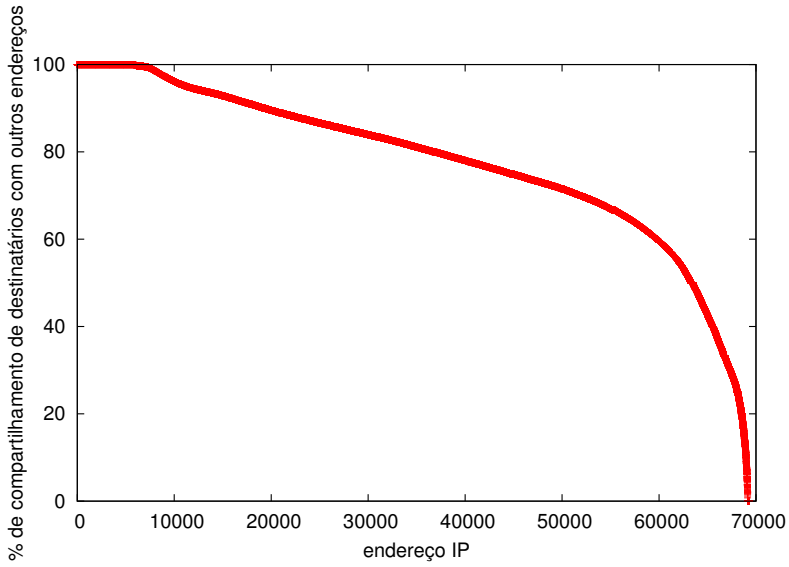

Figura 3. proporção de destinatários de cada endereço IP compartilhada por outros IPs

de endereços IP. Para analisar as sobreposições, montamos um grafo em que os nós são os endereços IP e as arestas entre dois endereços indica que dois endereços compartilham um número significativo de destinatários. Em termos de metodologia, a estratégia de aglutinar endereços IP a partir de um grafo foi apresentada no CEAS 2006 [Li and Hsieh 2006]. Os autores propõem o agrupamento de endereços IP a partir de critérios como a utilização de uma URL comum ou a menção a um valor monetário igual no corpo da mensagem. Nosso trabalho também agrupa endereços IP, porém, apoia-se em uma dimensão ainda não explorada para tal fim — os destinatários das mensagens.

Em nossos experimentos, verificamos que a sobreposição mínima de 5% é suficiente para agrupar as origens associadas ao mesmos destinatários. Exigimos que o compartilhamento ocorra nos dois sentidos. Caso contrário, um endereço IP que enviou poucas mensagens acabaria relacionado a um outro endereço que disparou um volume grande de mensagens, simplesmente porque os destinatários do primeiro acabariam ocorrendo entre o grande número de destinatários do segundo.

Essa estratégia é capaz de separar endereços IP que compartilham destinatários entre si de outros endereços que abusam outros destinatários, isolando e determinando diferentes agentes disseminadores de *spam*. Na próxima seção, detalhamos algumas propriedades dos agrupamentos de endereços IP encontrados.

4. Caracterização de *Spammers*

A instanciação do grafo de endereços IP para o conjunto de dados considerado no trabalho resultou na estrutura exibida na figura 4. É possível verificar que existem diversos *componentes* no grafo, representando diferentes grupos de endereços IP que consistentemente enviam *spams* a conjuntos similares de destinatários. Cada componente é bastante denso, possuindo arestas entre quase todos os seus membros. Alguns grupos, inclusive, chegam a formar um *clique*, um sub-grafo em que há ligações entre todos os seus nós.

As tabelas 2 e 3 exibem informações gerais sobre os grupos de endereços IP en-

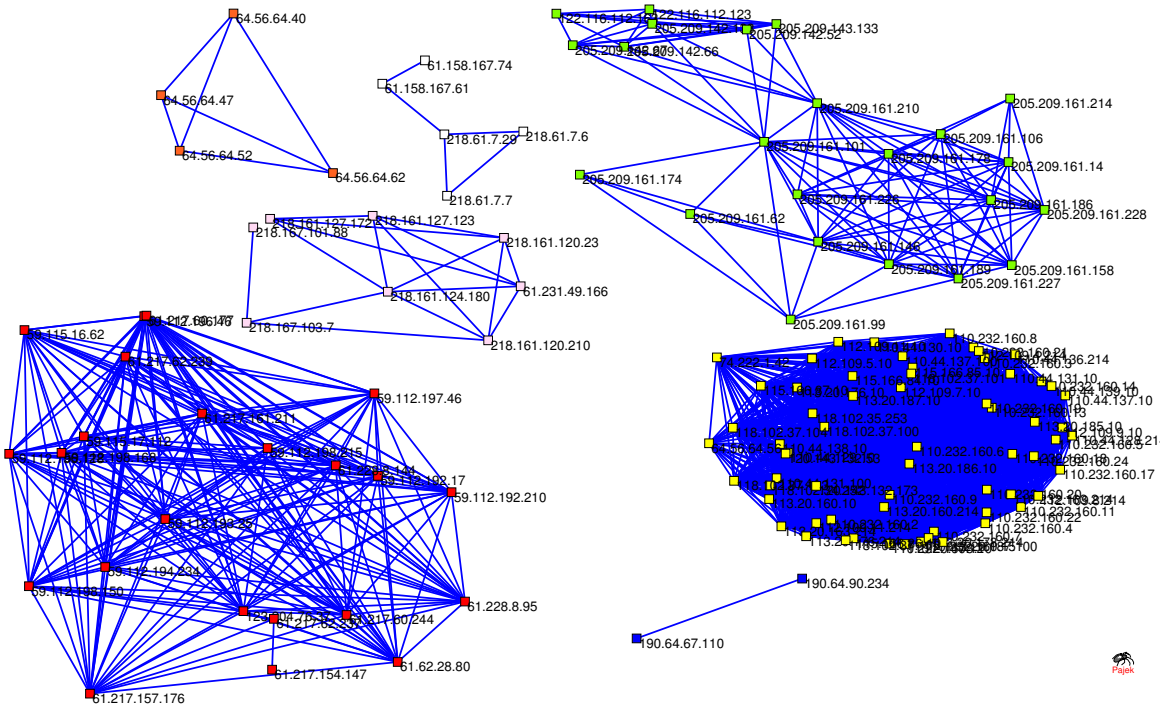

Figura 4. grafo relacionando endereços IP que compartilham pelo menos 5% de destinatários

contrados em relação à características de rede e de conteúdo, respectivamente. Os sete grupos de *spammers*, que respondem por cerca de 75% do tráfego de mensagens de *spam* observado no nosso conjunto de dados, estão associados a origens diversas: Ásia (Filipinas, Taiwan e China), Estados Unidos e Uruguai. É interessante observar, no entanto, que cada grupo vem de um conjunto pequeno de endereços IP e muitos desses endereços estão alocados ao mesmo prefixo de rede. Essa observação evidencia a *localidade espacial* da origem dos abusos, que já é aplicada, por exemplo, como critério de detecção de *spams* em sistemas baseados em reputação de endereços IP [Hao et al. 2009]. Esses sistemas consideram que endereços IP próximos a outros que já enviaram *spam* possuem uma probabilidade alta de também serem *spammers*.

Tabela 2. Sumário Geral dos 7 maiores grupos disseminadores de *spam*

id	num. dest.	num. dest. únicos	CC de Origem	num. IPs distintos	num. prefixos de rede
1	127,6 milhões	3,1 milhões	PH	64	8
2	92,3 milhões	13,0 milhões	US	23	2
3	69,6 milhões	2,6 milhões	TW	23	3
4	48 mil	42 mil	UY	2	1
5	591 mil	397 mil	TW	8	2
6	6,2 milhões	3,7 milhões	CN	5	2
7	186,6 milhões	15,7 milhões	US	4	1

A tabela 3 mostra objetivos e domínios de destino variados para as campanhas de *spam* (determinadas pela técnica apresentada em [Calais et al. 2008]). Os endereços IP do grupo 6, em particular, disseminaram 5 campanhas diferentes. Essas campanhas pode ser evoluções uma das outras ou representar um caso em que um *spammer* se concentra em disseminar campanhas de terceiros. Um trabalho futuro interessante é comparar essas

campanhas em termos de conteúdo e estratégia de construção. O grupo 7 dissemina um tipo de mensagem que não pretende vender nenhum produto ou serviço; são mensagens que contém puramente texto aleatório com o propósito de viciar filtros de *spam* baseados em aprendizagem estatística. Este seria um abuso difícil de detectar aplicando-se unicamente uma técnica de detecção de campanhas, pois há poucas características que se mantém invariantes nesse tipo de ataque.

Tabela 3. Conteúdo e Principal Domínio de Destino dos 7 maiores grupos disseminadores de *spam*

id	duração (dias)	num. campanhas	assunto	principal domínio de destino
1	82	1	venda de produtos	ezweb.ne.jp
2	101	1	fraude	yahoo.com.tw
3	45	1	?	msa.hinet.net
4	3	2	venda de mobílias / TVs	adinet.com.uy
5	10	1	venda de eletrônicos	yahoo.com.tw
6	72	5	produtos farmacêuticos	yahoo.com
7	29	1	ataque ao filtro	yahoo.com

Em relação à quantidade de destinatários de cada grupo, é possível observar, ainda na tabela 3, que o número de destinatários distintos em cada grupo é consideravelmente menor que o número total de destinatários, evidenciando que esses *spammers* iteram várias vezes sobre suas listas. Para entender melhor esse fenômeno, plotamos para cada grupo o número médio de observações anteriores de cada destinatário para cada dia em que o grupo enviou *spams*. O resultado é exibido na figura 5.

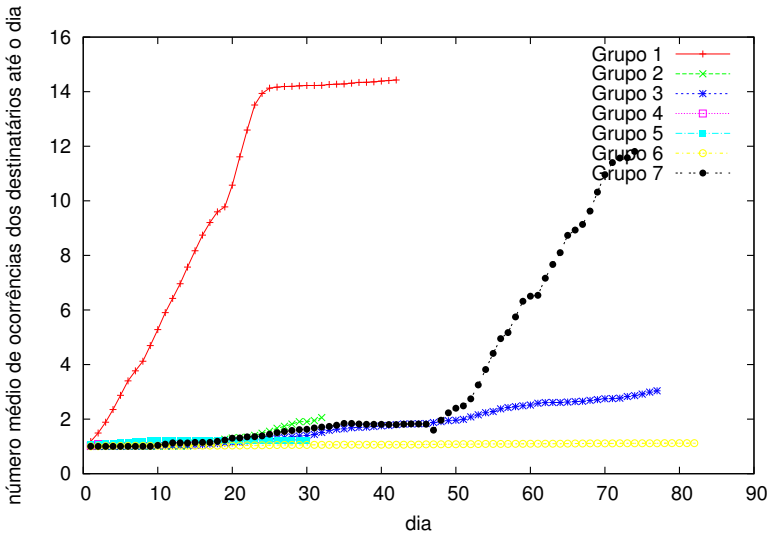

Figura 5. Número médio de ocorrências acumuladas até cada dia, para os destinatários de cada grupo

Podemos observar diferentes estratégias entre os grupos. Os grupos 1 e 7 parecem iterar várias vezes sobre suas listas. Após 30 dias de abuso, o grupo 1 entregou 14 mensagens a cada destinatário da sua lista, em média. A taxa de repetição cresce lentamente ao longo dos dias, evidenciando as iterações que o *spammer* executa sobre a sua lista. O comportamento do grupo 7 divide-se em duas fases distintas: uma em que poucos

destinatários se repetem, até os 50 primeiros dias de abuso, e outra em que a reincidência dos destinatários aumenta gradativamente. Essa curva pode indicar a mudança de estratégia do *spammer* na forma como ele distribui suas mensagens a seus destinatários ou um aumento significativo na taxa de envio de suas mensagens. Para os outros grupos, a taxa de reincidência de destinatários é praticamente estável, permanecendo entre 1 e 2 observações por todo o período de abuso, indicando que esses abusos não se estabilizaram e que os *spammers* associados a esses grupos ainda estão disseminando suas listas. Por outro lado, a tendência é que as listas dos grupos 1 e 7 já tenham sido observadas por completo, o que é coerente com a observação de que as listas associados a esse dois grupos são consideravelmente maiores que as outras (tabela 2). Os grupos 2, 3, 4, 5 e 6, provavelmente, correspondem a *spammers* que distribuem suas listas por muitas máquinas com vulnerabilidades, fora da visão dos nossos *honeypots*. A identificação de listas de destinatários, portanto, pode indicar o quão completa é a visão do comportamento de cada agente disseminador de *spam*, que, com qualquer mecanismo de coleta, sempre vai estar limitada a amostras. O problema da amostragem dos dados é um caso típico em que a determinação de listas de destinatários complementa o conhecimento adquirido a partir dos agrupamentos tradicionalmente determinados para caracterizar *spammers*, a partir de suas campanhas de *spam*.

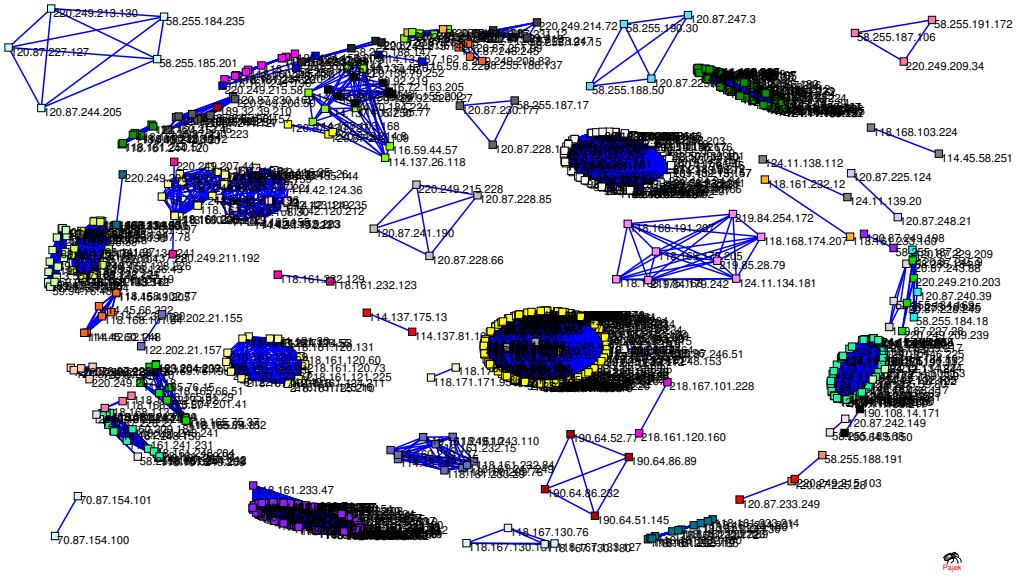

Figura 6. grafo relacionando endereços IP que compartilham 100% de destinatários

A fim de encontrar outros grupos de endereços IP potencialmente associados a um único *spammer*, verificamos quais os grupos obtidos ao ajustar o limiar de compartilhamento mínimo para 100%. O grafo resultante está exibido na figura 6. Novamente, diversos componentes são identificados, representando diferentes grupos de endereços IP que abusam exatamente o mesmo conjunto de destinatários. Em particular, esses grupos representam, em sua maioria, *spammers* que testaram o abuso ao *proxy* ou *mail relay* aberto emulados por um dos *honeypots* e enviaram mensagens unicamente para seu endereço de correio eletrônico de teste. A identificação desses grupos é importante, portanto, para avaliação e monitoração dos *spammers* que consideram abusar o *honeypot* e não o fazem, e podem ajudar a suportar a melhoria da arquitetura do *honeypot* para cap-

turar tráfego que não estaria sendo observado. Alguns desses grupos enviaram centenas de mensagens de teste durante o período de coleta, revelando a agressividade com que alguns *spammers* procuram máquinas com falhas de segurança disponíveis na Internet.

Por fim, analisamos como os *spammers* identificados repartem suas listas de destinatários verificando o caractere inicial dos endereços de correio alvo de suas mensagens. Primeiramente, edxtraímos de cada endereço de correio eletrônico o primeiro caractere e computamos a frequência de cada um no nosso conjunto de dados. Para cada endereço IP, computamos a sequência de caracteres iniciais presente em sua lista de endereços. Por exemplo, um endereço IP com a sequência [abrt] abusou endereços que começam com as iniciais *a*, *b*, *r* e *t*, independentemente da quantidade de mensagens enviada para cada inicial. A figura 7 exibe o tamanho médio dessas listas para as 30 sequências mais frequentes. O eixo Y está em escala logarítmica.

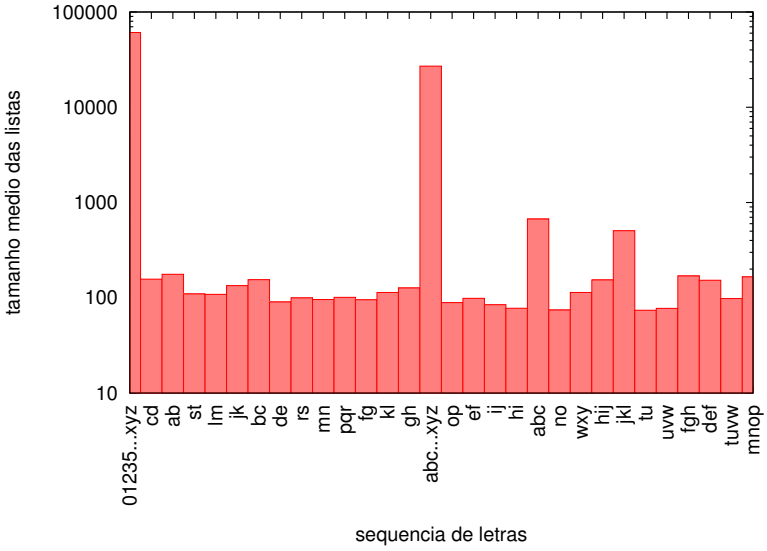

Figura 7. tamanho médio das listas para cada sequência de caractere inicial

O primeiro padrão que pode ser observado é que é muito comum *spammers* disseminarem suas campanhas dividindo a lista de destinatários alfabeticamente e distribuindo os destinatários entre diferentes máquinas. O histograma indica duas estratégias de disseminação de *spams* distintas: endereços IP que abusam destinatários em todo o espaço de caracteres iniciais em grande quantidade, sem distribuir os destinatários, e endereços que entregam poucas mensagens a destinatários cuja inicial é próxima ou igual àquela dos destinatários associados ao mesmo abuso. Clayton [Clayton 2008] demonstrou que o caractere inicial dos endereços de correio eletrônico pode afetar o volume de *spam* recebido e que muitos *spammers* percorrem suas listas de destinatários em ordem alfabética. As inicias que mais recebem *spams* apontadas pelo trabalho incluem *a*, *b*, *c*, *j*, *k* e *l*, que estão presentes em dois picos na figura 7.

Esse padrão de distribuição de listas reforça um padrão observado em um trabalho anterior [Calais et al. 2008] de que endereços IP que enviam *spams* em grande quantidade em geral estão associados à origem real do *spammer*, e endereços IP que enviam poucos *spams* em geral são máquinas na rede, como *bots* ou *relays* abertos, que entregam apenas porções da lista de destinatários do *spammer*.

5. Trabalhos Relacionados

A literatura aponta alguns trabalhos que investigam, de alguma forma, os destinatários de mensagens de *spam*. Esses trabalhos podem ser divididos em duas categorias: aqueles que verificam como as listas de destinatários são obtidas e os que investigam como elas são distribuídas.

A maior parte dos trabalhos relacionados a listas de destinatários de *spammers* recai na primeira categoria. Os resultados dessas análises indicam que um endereço de correio eletrônico postado em uma página na Web normalmente é rapidamente encontrado por algum *spammer* (muitas vezes, após algumas horas de exposição), o que reflete a maneira sistemática e agressiva com a qual *spammers* procuram endereços para compor e aumentar suas listas de alvos [Prince et al. 2005]. Alguns perfis de coletores de endereços também foram determinados, como aqueles que demoram a enviar *spams* para os alvos depois que eles são coletados, mas o fazem em grande quantidade disseminando *spams* de venda de produtos, e aqueles que enviam mensagens para os endereços logo depois de coletá-los, entregando especialmente mensagens de fraude.

Neste trabalho, não verificamos a forma como os endereços são capturados, mas buscamos encontrar agrupamentos entre as origens de *spam* a partir de padrões de tentativas de envio a esses destinatários na rede. Nesse sentido, já foi determinado que a relação entre origens e destino das mensagens é diferente entre os *spams* e mensagens legítimas [Gomes et al. 2004]. No caso das mensagens legítimas, o fluxo de mensagens é bilateral, ou seja, destinatários recebem mensagens e respondem ao emissor. No caso dos *spams*, a comunicação é estritamente unilateral.

Em relação à entrega dos *spams* aos alvos, existem trabalhos que analisam o sistema de entrega de *spams* de *botnets* e detalham a estratégia de disseminação de listas por meio delas [John et al. 2009, Kreibich et al. 2008]. Esses trabalhos estimam o tamanho das listas associadas a cada *botnet* em centenas de milhões de destinatários e mostram que a sobreposição entre listas de *botnets* diferentes não é muito alta, variando entre 6% e 28% [John et al. 2009]. Esse resultado é um bom indicador de que destinatários podem ser úteis para identificar *spammers*, motivação principal deste artigo.

6. Conclusão e Trabalhos Futuros

Neste trabalho, mostramos que listas de destinatários fornecem um insumo interessante para se estudar o comportamento de *spammers* e que a repetição de destinatários pode ser utilizada como critério de identificação e caracterização de grupos de endereços que agem coordenadamente, abusando a mesma lista de destinatários. Nossa proposta consiste em isolar spammers por meio da reconstrução das suas listas de destinatários, para então estudar o comportamento de cada um na rede. A filtragem de spams não é nosso objetivo imediato neste trabalho.

O principal objetivo do trabalho é mostrar que os destinatários dos *spams* – dimensão regularmente ignorada em estudos de caracterização de spammers – e a reconstrução das listas de e-mail podem ser explorados como instrumento para caracterizar e entender como *spammers* atuam na rede e disseminam o conteúdo dos seus spams. Apesar de sabermos que spammers trabalham com listas de endereços, não encontramos registros científicos desse fato, muito menos uma caracterização da forma como as mesmas são utilizadas.

A identificação dos *spammers* é feita a partir da instanciação de um grafo que contempla sobreposições entre as listas de destinatários de cada origem de *spam* e os componentes do grafo determinam diferentes *spammers*.

A metodologia foi aplicada a um conjunto de dados real e mostrou que, para um universo de 40 milhões de mensagens, menos de uma dezena de *spammers* é responsável por um grande volume do tráfego (75%). Entre os padrões de comportamento descobertos, mostramos que os principais *spammers* encontrados enviam mensagens de um conjunto pequeno de endereços IP, alocados a poucos prefixos de rede diferentes. Embora os resultados não possam ser generalizados, pois se baseam em um conjunto de dados específico, demonstram o potencial da estratégia.

Acreditamos que o trabalho abre novas possibilidades de análise do comportamento dinâmico e evolutivo de *spammers*. A partir da identificação das entidades geradoras de *spam*, é possível observar as várias campanhas associadas a cada um e comparálas. Podemos, ainda, combinar diferentes dimensões associadas aos *spams*, como sua origem, conteúdo e destino (dimensão explorada neste trabalho) em um arcabouço único de identificação e caracterização de *spams*.

Uma outra linha de pesquisa consiste em analisar a dinâmica do grafo de endereços IP, verificando em que momentos as arestas surgem e se componentes e partições detectados são estáticos ou mudam ao longo do tempo.

7. Agradecimentos

O presente trabalho foi realizado com o apoio do UOL (`www.uol.com.br`) , através do Programa UOL Bolsa Pesquisa, Processo Número 20090215215700 e do NIC.br.

Referências

Calais, P. H., Guedes, D., Jr., W. M., Hoepers, C., and Steding-Jessen, K. (2008). Caracterização de estratégias de disseminação de spams. In *Anais do Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos*, Rio de Janeiro, RJ.

Calais, P. H., Guedes, D., Wagner Meira, J., Hoepers, C., Chaves, M. H. P. C., and Steding-Jessen, K. (2009). Spamming chains: A new way of understanding spammer behavior. In *Proceedings of the 6th Conference on e-mail and anti-spam (CEAS)*, Mountain View, CA.

Clayton, R. (2008). Do zebras get more spam than aardvarks? *The 5th Conference on Email and Anti-Spam (CEAS) 2008*.

Cranor, L. F. and LaMacchia, B. A. (1998). Spam! *Commun. ACM*, 41(8):74–83.

Gomes, L. H., Cazita, C., Almeida, J. M., Almeida, V., and Meira, Jr., W. (2004). Characterizing a spam traffic. In *IMC '04: Proceedings of the 4th ACM SIGCOMM conference on Internet measurement*, pages 356–369, New York, NY, USA. ACM.

Goodman, J., Cormack, G. V., and Heckerman, D. (2007). Spam and the ongoing battle for the inbox. *Commun. ACM*, 50(2):24–33.

Hao, S., Syed, N. A., Feamster, N., Gray, A., and Krasser, S. (2009). Detecting Spammers with SNARE: Spatio-temporal Network-level Automatic Reputation Engine. In *Usenix Security '09*, Montreal, Canada.

Hayes, B. (2003). Spam, spam, spam, lovely spam. *American Scientist*, 91(3):200–204.

John, J. P., Moshchuk, A., Gribble, S. D., and Krishnamurthy, A. (2009). Studying spamming botnets using botlab. In *NSDI'09: Proceedings of the 6th USENIX symposium on Networked systems design and implementation*, pages 291–306, Berkeley, CA, USA. USENIX Association.

Kokkodis, M. and Faloutsos, M. (2009). Spamming botnets: Are we losing the war? *The 6th Conference on Email and Anti-Spam (CEAS) 2009*.

Krawetz, N. (2004). Anti-honeypot technology. *IEEE Security & Privacy*, 2(1):76–79.

Kreibich, C., Kanich, C., Levchenko, K., Enright, B., Voelker, G. M., Paxson, V., and Savage, S. (2008). On the spam campaign trail. In *LEET'08: Proceedings of the 1st Usenix Workshop on Large-Scale Exploits and Emergent Threats*, pages 1–9, Berkeley, CA, USA. USENIX Association.

Li, F. and Hsieh, M.-H. (2006). An empirical study of clustering behavior of spammers and group-based anti-spam strategies. *Proceedings of the Third Conference on Email and Anti-Spam (CEAS). Mountain View, CA*.

MAAWG (2009). Email Metrics Program: Report #5 – Third and Fourth Quarter 2008. `http://www.maawg.org/about/MAAWG_2008-Q3Q4_Metrics_Report.pdf`.

Pathak, A., Hu, Y. C., and Mao, Z. M. (2008). Peeking into spammer behavior from a unique vantage point. In *LEET'08: Proceedings of the 1st Usenix Workshop on Large-Scale Exploits and Emergent Threats*, pages 1–9, Berkeley, CA, USA. USENIX Association.

Prince, M. B., Holloway, L., Langheinrich, E., Dahl, B. M., and Keller, A. M. (2005). Understanding how spammers steal your e-mail address: An analysis of the first six months of data from project honey pot. *Proceedings from CEAS'05: Conference on Email and Anti-Spam.*

Provos, N. and Holz, T. (2007). *Virtual Honeypots: From Botnet Tracking to Intrusion Detection*. Addison-Wesley Professional, 1st edition. ISBN-13: 978-0321336323.

Pu, C. and Webb, S. (2006). Observed trends in spam construction techniques: a case study of spam evolution. *Proceedings of the 3rd Conference on Email and Anti-Spam (CEAS)*.

Sipior, J. C., Ward, B. T., and Bonner, P. G. (2004). Should spam be on the menu? *Commun. ACM*, 47(6):59–63.

Xie, Y., Yu, F., Achan, K., Gillum, E., Goldszmidt, M., and Wobber, T. (2007). How dynamic are ip addresses? In *SIGCOMM '07: Proceedings of the 2007 conference on Applications, technologies, architectures, and protocols for computer communications*, pages 301–312, New York, NY, USA. ACM.

Yeh, C.-C. and Lin, C.-H. (2006). Near-duplicate mail detection based on url information for spam filtering. In *Information Networking. Advances in Data Communications and Wireless Networks*, pages 842–851. Springer Berlin / Heidelberg.