Uso de Honeypots de Baixa Interatividade para o Estudo do Abuso de Proxies Abertos para o Envio de Spam

Klaus Steding-Jessen[1], **Nandamudi L. Vijaykumar**[2], **Antonio Montes**[3]

[1]Núcleo de Informação e Coordenação do Ponto br - NIC.br
Centro de Estudos, Resposta e Tratamento de Incidentes de Segurança - CERT.br
São Paulo (SP)

[2]Instituto Nacional de Pesquisas Espaciais - INPE
Laboratório Associado de Computação e Matemática Aplicada - LAC
São José dos Campos (SP)

[3]Ministério da Ciência e Tecnologia - MCT
Centro de Pesquisas Renato Archer - CenPRA
Campinas (SP)

jessen@cert.br, vijay@lac.inpe.br, antonio.montes@cenpra.gov.br

Resumo. *Um dos principais desafios para a mitigação do* spam *é a ausência de dados mais precisos. Este artigo descreve o projeto e implementação de uma arquitetura, com base em* honeypots*, para o estudo do* spam *e sua mitigação. Também são discutidos os resultados obtidos a partir de 15 meses de coleta.*

Resumo. *One of the main challenges to mitigate the spam problem is the lack of precise information about it. This paper describes the design and implementation of a honeypot based architecture built to study spam and ways to mitigate it. The results based on 15 months of captured spams are also discussed.*

1. Introdução

O *spam* é uma das formas de abuso da Internet que mais tem crescido, atualmente sendo responsável por uma parcela significativa dos *emails* que trafegam na rede [Hayes 2003]. Além disso, o *spam* tem sido amplamente utilizado para enviar mensagens relacionadas com *phishing* (mensagem que procura induzir usuários ao fornecimento de dados pessoais e financeiros) e para disseminação de códigos maliciosos [Milletary 2005].

Um dos problemas atuais para a criação de mecanismos efetivos para a mitigação do problema é a ausência de dados mais precisos sobre a sua dimensão e sobre os mecanismos utilizados para a disseminação de *spam* e códigos maliciosos.

Este artigo descreve o projeto e implementação de uma arquitetura, com base em *honeypots*, para o estudo do problema do *spam*, em particular:

- abuso de *relays* e *proxies* abertos, tradicionalmente usados para o envio de *spam* e outras atividades maliciosas [Hoepers et al. 2003, Krawetz 2004];

- abuso de máquinas de usuários finais, conectados via banda larga, que estão vulneráveis e tem o potencial de serem infectadas por códigos maliciosos e controladas remotamente para o envio de *spam* [Sauver 2005, Holz 2005].

A arquitetura implementada é composta de um conjunto de sensores, baseado em *honeypots*, para a captura de *spam* e posterior coleta por um servidor centralizado. Este artigo também descreve as estensões implementadas na funcionalidade do *software* `Honeyd` de modo a aumentar a eficiência da captura de *spam*.

2. Rede de Honeypots para Captura de Spam

Spammers continuamente realizam varreduras na Internet por computadores com *proxies* abertos [Krawetz 2004]. Uma vez localizados, estes computadores são então explorados para efetuar conexões para os servidores SMTP dos destinatários do *spam*. Na arquitetura implementada os *honeypots* foram colocados de modo a simular computadores com *proxies* abertos. Desse modo, um *spammer* que tentar abusar de um destes honeypots para o envio de *spam*, será levado a acreditar está tendo sucesso em enviar seus *emails*.

Esta arquitetura contou com 10 *honeypots* de baixa interatividade, responsáveis pela coleta de *spams*, instalados em redes de banda larga de 5 operadoras diferentes (cabo e ADSL). Os *honeypots* foram instalados nas residências de voluntários do projeto para refletir as condições a que estão submetidos computadores típicos de usuários residenciais com conexões banda larga.

3. Resultados

Tabela 1. Estatísticas Gerais.

Início da coleta dos dados	10/06/2006
Fim da coleta dos dados	18/09/2007
Dias de dados coletados	466
Total de emails coletados	524.585.779
Total de destinatários	4.805.521.964
Média de destinatários por spam	9,16
Média de emails por dia	1.125.720,56
IPs únicos que enviaram spam	216.888
ASNs únicos que enviaram spam	3.006
Países (country codes) de origem	165

Durante 15 meses de operação, foram coletados mais de quinhentos milhões de *spams* que seriam entregues a mais de 4 bilhões de destinatários, como mostrado na Tab. 1. Nesta seção são mostrados com mais detalhes alguns resultados preliminares da análise desses *emails*.

Na Tab. 2 é possível ver uma grande concentração na origem dos *emails*, com apenas três países sendo responsáveis por mais de 90% de todas as mensagens.

4. Conclusões

A análise dos *spams* recebidos deixou claro que os endereços IPs que exploram *proxies* abertos para o envio de *spam* estão extremamente concentrados: aproximadamente 98% de todos os *emails* vieram de 10 países diferentes apenas. Aproximadamente 90% das mensagens vieram de dois países: Taiwan e China.

Tabela 2. *Country Codes* (CC) mais freqüentes.

#	CC	*Emails*	%
01	TW	385.189.756	73,43
02	CN	82.884.642	15,80
03	US	29.764.293	5,67
04	CA	6.684.667	1,27
05	JP	5.381.192	1,03
06	HK	4.383.999	0,84
07	KR	4.093.365	0,78
08	UA	1.806.210	0,34
09	DE	934.417	0,18
10	BR	863.657	0,16

Os dados foram apenas analisados em função da sua origem, não do seu conteúdo. Como um trabalho futuro é possível a análise mais detalhada dos *emails* coletados, observando-se o corpo das mensagens, endereços de destinatários, entre outros.

Referências

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

[Hoepers et al. 2003] Hoepers, C., Steding-Jessen, K., and Chaves, M. H. P. C. (2003). Projeto e Desenvolvimento de um Sistema de Controle e Acompanhamento de Notificações de Spam. In *Anais do V Simpósio sobre Segurança em Informática (SSI'2003)*, São José dos Campos, SP.

[Holz 2005] Holz, T. (2005). A short visit to the bot zoo. *IEEE Security & Privacy*, 3(3):76–79. `http://www-pi1.informatik.uni-mannheim.de/publications/show/13`.

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

[Milletary 2005] Milletary, J. (2005). Technical trends in phishing attacks. `http://www.cert.org/archive/pdf/Phishing_trends.pdf`. CERT Coordination Center, Carnegie Mellon University.

[Sauver 2005] Sauver, J. S. (2005). Spam zombies and inbound flows to compromised customer systems. In *Procedings of the MAAWG General Meeting*. `http://darkwing.uoregon.edu/~joe/zombies.pdf`.