GOOGLE HACKING

Alex Sander de Oliveira Toledo[1]
Dayene Ângela de Almeida[2]

RESUMO: O presente trabalho apresenta os recursos que o Google oferece no que diz respeito à coleta de informações. Estes recursos servirão de referência aos usuários web quanto à segurança, na espera de bloquear formas de vazamento de informações.

PALAVRAS-CHAVE: Informações, Segurança, Google Hacking.

1 Introdução

Existem ferramentas que são de grande valor e na maioria das vezes não damos tanta importância pela sua aparência, normalmente por apresentarem ser o que qualificamos como simples. E por muitas das vezes o "simples" aspecto na internet pode ser um grande problema de segurança.

A segurança na internet é um tema critico nos dias atuais. A cada dia surgem diversas formas de testes para evitar invasões e ataques às informações das organizações e em sistemas pessoais o que ajuda evitar os prejuízos financeiros, perdas de informações confidenciais ou importantes. Por isso a segurança da informação protege e mantem as informações sigilosas. (PELTIER, 2001).

A internet pode esconder intenções maliciosas e deixar enganar pela inocência visual. Isto sugere que qualquer coisa que não está visível não pode trazer insegurança, como é o caso da visão da maioria dos usuários quanto ao serviço de busca: Google. Segundo Felipini (2008), o Google é um dos serviços de busca mais utilizado na internet, parecendo ser inofensivo. Por de traz da aparência simples ele esconde uma poderosa ferramenta para busca de informações de todos os tipos.

O uso malicioso pode ser feito através de uma pirataria informatizada, Google Hacking, que utiliza o Google para encontrar falhas de segurança em configurações, sites e códigos de computador.

A arte do Google Hacking utiliza o Google para localizar informações confidenciais ou para encontrar vulnerabilidade que podem ser exploradas, podendo ser uma grande ameaça. Segundo Pinheiro (2010), o que explica essas ameaças é a utilização do agente de busca para catalogar paginas, ao mesmo tempo em que aumenta de forma considerável o numero de respostas da ferramenta, traz o problema de muitas vezes catalogar mais do que os administradores de paginas gostariam, porque os muitos servidores estão conectados a Internet.

2 Falhas Reconhecidas

O motor de busca Google, é fácil de usar (figura 1), encontrado em www.google.com,oferece benefícios óbvios para qualquer tipo de internauta: leigos ou maliciosos. E para uso malicioso usa- se a técnica de Google Hacking.

Figura 1: A página de pesquisa do Google.
Fonte:http://www.google.com.br

A barra (campo) de pesquisa permite ao usuário digitar um termo de pesquisa ou uma palavra como, por exemplo, "jornal" e teremos muitas opções disponíveis. Porem o Google não oferece apenas esse campo, há vários campos como, por exemplo, "Ferramentas de Idiomas" que permite ao usuário definir opções diferentes de idiomas e outros campos que podem ser de grande ajuda.

O Google usa operadores booleanos para restringir suas pesquisas. Segundo Aamato (2005) os operadores booleanos são palavras que informam ao sistema como combinar os termos de sua pesquisa. São eles: AND, OR, NOT, respectivamente, E, Ou e NÃO. Ao fornecermos uma lista de termos de busca, o Google automaticamente tenta encontrar todas as palavras em uma lista, tornando o operador booleano AND redundante.

O Google é fácil de ser usado, não há necessidade de descrever a sua funcionalidade básica. Em vez disso, vamos conhecer os operadores disponíveis:

a)O sinal de mais (+) é usado com uma maneira de substituir um valor booleano AND. Mas o Google usa o sinal de mais (+) de uma forma diferente. Usa para forçar uma busca por uma palavra muito comum.

b)O sinal de menos (-) para excluir um termo de uma busca. Lembrando que não pode haver espaço de um sinal com a busca, porque pode produzir resultados diferentes.

c)As aspas duplas (" "), ou busca mista pode envolver ambas as frases e termos individuais. Exemplos: "Microsoft Office", o que resultara apenas o que inclui a frase.

d)Um ponto (.) serve como um único caractere.

e)Um asterisco (*), ou curinga, diz ao Google para tentar tratar o (*) como um espaço reservado para qualquer termo desconhecido e depois encontrar os melhores resultados. O Google permite o uso de operadores para ajudar a filtrar as pesquisas e o uso avançado do mesmo é muito simples.

f)Operador: SEARCH_TERM. Palavra chave.

Alguns operadores podem ser usados como uma consulta independente. Dessa forma, muitas informações "apuradas" podem ser encontradas, basta saber o que procurar. O ainda armazena no "cache" uma versão de paginas e arquivos, de modo que mesmo corrigindo o problema no servidor, os dados podem ainda estar expostos. Alguns tipos de ataques que podem ser feitos através do Google:

a)Site: encontra as paginas da Web em um site ou domínio especifico. Exemplo: site: sistemas de informação newtonpaiva.br. O resultado mostra sobre o curso de Sistemas de Informação no site da Newton Paiva.

b)Filetype: operador instrui o Google a pesquisar somente dentro do texto de um determinado tipo de arquivo. Exemplo: filetype:txt coração. Os tipos de arquivos mais comuns são:

Adobe Flash (.swf);
Adobe Portable Document Format (.pdf);
Adobe PostScript (.ps);
Autodesk Design Web Format (.dwf);
Google Earth (.kml, .kmz);
GPS eXchange Format (.gpx);
Hancom Hanword (.hwp);
HTML (.htm, .html, other file extensions);
Microsoft Excel (.xls, .xlsx);
Microsoft PowerPoint (.ppt, .pptx);
Microsoft Word (.doc, .docx);
OpenOffice presentation (.odp);
OpenOffice spreadsheet (.ods);
OpenOffice text (.odt);
Rich Text Format (.rtf, .wri);
Scalable Vector Graphics (.svg);
Text (.txt, .text, other file extensions), incluindo o código fonte em linguagem de programação comum:
Basic source code (.bas);
C/C++ source code (.c, .cc, .cpp, .cxx, .h, .hpp);
C# source code (.cs);
Java source code (.java);
Perl source code (.pl);
Python source code (.py);
Wireless Markup Language (.wml, .wap);
XML (.xml);

c)Link: o Google pesquisa dentro de um hyperlink para uma busca. Exemplo: link: www.apple.com/iphone.

d)Cache: exibe a versão de uma pagina web quando o Google busca no site. A URL do site deve ser fornecida após os dois pontos. Exemplo: cache:www.newtonpaiva.br. Este é o cache do Google de http://www.newtonpaiva.br/.

e)Intitle: para pesquisar um termo no titulo de um documento. Exemplo: intitle:index.of. mp3 Fernando e Sorocaba.

f)Inurl: para pesquisar somente dentro da URL (endereço web) de um documento. Exemplo: inurl:musica.

3 Como Realizar

Como foi mostrado no tópico anterior existem operadores simples que ajudam a revirar ou incrementar a busca. É essa simplicidade, junto ao poder dos resultados, que transformam o Google em uma ideia tão sensacional.

O hacking utiliza a Barra de Ferramenta do Google ou Google Toolbar que é instalado no browser por meio de um executável, a descobrir informações, descobrir sobre o assunto ou identidades de pessoas e quebra barreiras de segurança usando uma pesqui-

sa no Google inocente e sem malicia, como dissimulação.

O Google é apenas um instrumento utilizado para explorar falhas do sistema operacional, do planejamento de TI ou do próprio administrador de sistemas, então o que mostra que ele não tem responsabilidade direta sobre a vulnerabilidade. Com ajuda do FrontPage, o software de manufatura e editoração de paginas Web da Microsoft, empresas sem administradores de segurança ou pequenas lojas online também podem causar sua própria falência, com um auxilio mínimo dos Hackings utilizadores do Google.

Um hacker procura uma informação pelo termo:

Inurl:"ordes.txt"

Figura 2: Usando o operador Inurl:"ordes.txt"
Fonte: http://www.eecoswitch.com/f2m/orders.txt

O resultado é brilhante (figura 2), são arquivos de mais de 2MB, reproduzindo bancos de dados com inúmeras informações como numero de cartão de credito, endereço, telefone e senhas de validação de compras.

Já usando:

inurl:"auth_user_file.txt"

Figura 3: Usando o operador inurl:"auth_user_file.txt".
Fonte: http://www.asklandskoca.com/cgi-local/Copyforum/User_info/admin.session

Com esse comando, é fácil encontrar arquivos (figura 3) contendo listas de usuários, senhas e diretórios de login.

A combinação:
Index of /admin

Figura 4: Usando o operador Index of /admin.
Fonte: [illegible]

Equivale a entrar no mundo dos servidores desprotegidos e colocados no ar em um dia por equipes apresentadas. E uma busca rápida retorna nada mais do que 3600 resultados (figura 4).

A expressão:
inurl:”form_results.txt”

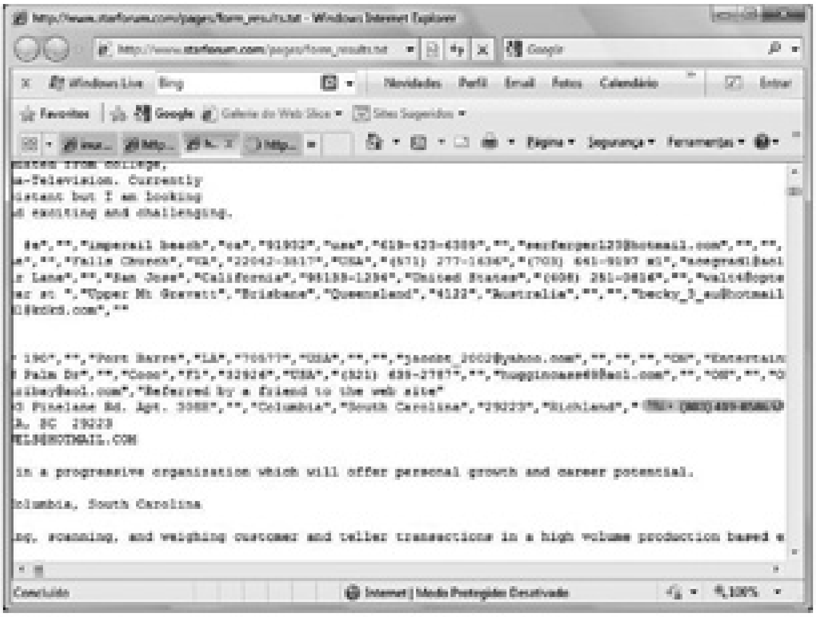

Figura 5: Usando o operador inurl:"form_results.txt".
Fonte: http://www.starforum.com/pages/form_results.txt.

Encontramos vários sites expostos (figura 5). Alguns mostram nomes, números de telefone e endereço de email dos seus visitantes.

Tudo o que vimos mostra o grande poder do Google e como é fácil descobrir dados e informações. Para ficar ainda mais claro o uso dessas operações através do Google veja o exemplo: imagine que um hacker invada um site que possua uma media quantidade de visitantes, consegue invadir uma conta de administrador. Com a senha de administrador qualquer um pode fazer um upload de arquivos executáveis, escondendo-os as arvores e diretórios, sem definir links definidos. E se o verdadeiro administrador não tiver o costume de olhar sua pagina não notará a inclusão de links com códigos maliciosos. Ainda há muitas vulnerabilidades que podem e são explorados. Basta ter um hacker, um usuário ou servidor, e um browser rodando no Google.

Segundo SOUSA e ROCHA (2010) algumas grandes empresas de tecnologia como o Google, por exemplo, investem no desenvolvimento dos seus próprios SGBDs baseado na ideia NoSQL[3] que ativa o processo de distribuição de arquivos ou bancos de dados usando funções de mapeamento e redução. E o Google desde 2004 investe, no Big Table, um banco de dados proprietário, desenvolvido para suprir as necessidades de armazenamento, totalmente baseado na filosofia de alto desempenho, escalabilidade e disponibilidade.

Portanto não podemos esquecer que nada que estiver on-line, mesmo que esteja escondido em varias camadas de subdiretórios, passa despercebido pelo Google. Segundo Gomes (2009) mesmo bancos ocultos por trás de proxy ou firewalls mal configurados podem ser encontrados, já que permitem que uma base de dados receba tentativas de requisição de todas as partes, sem restrição de IP, equivale a tornar o servidor de domínio publico. Ao haver duvida com relação à localização do endereço, basta o hacker utilizar ferramentas simples como o "tracert", uma ferramenta de redes do Windows que funciona em linha de comando, semelhante do "traceroute" do UNIX.

4 Como Prevenir

O sistema de busca do Google funciona bem demais, tudo que estiver on-line, mesmo que esteja escondido em varias camadas de subdiretórios, não passa despercebido pelo Google. Os hackers que se utilizam do Google ou de outros mecanismos de busca à procura de informações pessoais estão em busca de documentos do Word, bancos de dados, todas essas modalidades de arquivos possuem algum tipo de informação relevante para encontrar, por exemplo, planilhas de Excel, associando-as aos termos cartão, crédito e despesas ou um FTP (Protocolo de Transferência de Arquivos) totalmente aberto e cheio de informações.

Como já foi dito não podemos esquecer que a culpa dessas vulnerabilidades não são do Google. Ele é apenas um instrumento utilizado para explorar falhas do sistema operacional, do planejamento de TI ou próprio administrador de sistemas.

Então pra proteger as informações temos que manter os dados fora da Web. Mesmo inserindo dados temporários, existem grandes chances de esquecê-los ou que um Web Crawler, que é usado para manter uma base de dados atualizados, pode encontrá-los.

O melhor é usar formas mais seguras de compartilhamento de dados sensíveis como, por exemplo, SSH / SCP[4], os dados são cifrados durante a transferência, protege o usuário de invasões na rede evitando roubos de pacotes, senhas e dados diversos.

Usar as técnicas descritas nos tópicos acima pode ajudar a verificar o seu próprio site para informações sensíveis, dados ou arquivos vulneráveis.

Pode-se usar o Firewall como um sistema de controle aos acessos às redes de computadores que tem como objetivo permitir que somente pessoas autorizadas conectem a uma rede local ou privada de uma instituição, tornando uma excelente ferramenta de controle e uso da rede e da Internet, independendo do tipo de computador, seja doméstico e ou de servidor.

Segundo Alecrim (2004) um Firewall pode ser definido como uma barreira de proteção, que controla o tráfego de dados entre seu computador e a Internet (ou entre a rede onde seu computador está instalado e a Internet). Automaticamente o usuário terá como realizar transmissão e ou recebimento de dados autorizados e o funcionamento terá variação de acordo com o sistema e ou de como foi elaborado o programa. E uma rede com informações criptografadas em casos de transações de dados importantes podem também ajudar a proteger de ataques inesperados.

Fazer backup é importante para a segurança da informação. E pode ate mesmo fazer uma copia off-line dos dados mais importantes. Mas não adianta retornar com uma copia do backup se a vulnerabilidade explorada na invasão não for sanada.

Tem também o Google webmaster que é uma ferramenta desenvolvida pelo Google que proporciona aos usuários a obtenção de dados sobre rastreamento, indexação e tráfego de pesquisa e receber notificações sobre problemas no seu site. Todo o processo é simples e pode ser realizado através dos passos descritos na página: http://www.google.com/support/webmasters/bin/answer.py?answer=164734.Como existem pessoas capazes de explorar falhas na segurança das informações, a melhor prevenção contra vazamento é evitar vulnerabilidade, por isso o webmaster é uma ferramenta importante.

Portanto prevenir as informações contra problemas nada

mais é que agir contra agentes capazes de explorar falhas de segurança, evitando a vulnerabilidade das informações e dados.

5 Conclusão

A internet por si é uma ferramenta valiosa que nos ajuda a interagir com diversos ambientes, e o Google é um serviço on-line e software que hospeda e desenvolve uma series de serviços e produtos baseados na internet. O Google pode ser usado por usuários maliciosos a fim de atacar suas informações e pode ser utilizado para proteger as mesmas através de medidas simples de segurança e constante vigilância.

A utilização de operadores para catalogar páginas ao mesmo tempo em que aumenta, de forma ainda mais ampla, o numero de respostas do Google surpreende muitos administradores por perceberem que seus dados e informações não foram configurados de forma segura. E o Google Hacking encontra arquivos sigilosos ao fazer a busca através de operadores, expondo servidores na base de dados do Google na primeira oportunidade, já que não são capazes de diferenciar o que é publico do que é privado/confidencial.

Portanto o Google é indispensável quando se trata de buscas na web. E pode se concluir que Google Hacking ensina a descobrir, de forma discreta, segredos sobre a vulnerabilidade dos sites, como acessar bancos de dados, rastreamento de servidores, acessos a senhas secretas, etc., mas também nos ensina como se defender dos ataques digitais e evitá-los, ao revelar as técnicas de ataques na rede. E mesmo você não sendo um hacker há certas vantagens em conhecer a existência de hacking no Google porque, por exemplo, você pode descobrir informações confidenciais sobre você ou sua empresa e com isso pode ainda se proteger de eventuais vulnerabilidades, o que ajuda na proteção contra as ameaças de hacks do Google.

REFERÊNCIAS BIBLIOGRÁFICAS

AAMATO, Asclepios. **Operadores lógicos booleanos.** “AND”, ”OR” e ”NOt”. Disponível em:< http://www.asclepios.com.br/medico/content/operadores-l%C3%B3gicos-booleanos-%E2%80%9Cand%E2%80%9D-%E2%80%9Cor%E2%80%9D-e-%E2%80%9Cnot%E2%80%9D_>. Acesso: 22 set 2011.

ALECRIM, Emerson. **Firewall: Conceito e Tipos.** Disponível: <http://www.inforwester.com/firewall.php> . Acesso: 03 Nov 2011.

Felipini, Dailton. Top 10. Como alcançar o topo nos sites de busca. Disponível em: < http://www.e-commerce.org.br/sites_de_busca.php >. Acesso em: 15 set 2011.

GRIS, Grupo de resposta a incidentes de segurança. **Usando o Google como ferramenta hacker.** Google Hacking. Disponível em: <http://www.backtrack-linux.org/forums/tutoriais-e-howtos/26974-usando-o-google-como-ferramenta-hacker-%96-parte-1-google-hacking.html> . Acesso: 16 set 2011.

LONG, Johnny. **Google Hacking.** Mini Guide. Disponível em: < http://www.informit.com/articles/article.aspx?p=170880 >.Acesso 19 set 2011.

MUTHU, Sudar. **Operadores Avançados Google.** Como utilizar o Google com eficiência. Disponível em: < http://hackerproject.webnode.com.br/tutoriais/operadores%20avan%C3%A7ados%20google/ >. Acesso: 16 set 2011

Portal 5S Tecnologia. **Gerenciamento estratégico de segurança da informação.** Disponível em: < http://www.ietec.com.br/site/techoje/categoria/detalhe_artigo/261_>. Acesso em: 15 set 2011.

RAMOS, Robson. **O que é Google Hacking?** O que o Google faz e você não sabe!. Disponível em: <http://brainstormdeti.wordpress.com/2010/08/31/o-que-e-google-hacking-o-que-o-google-faz-e-voce-nao-sabe/> . Acesso em: 16 set 2011.

NOTAS DE RODEPÉ

1 Coordenador e Professor do Curso de Bacharelado em Sistemas de Informação do Centro Universitário Newton Paiva

2 Graduanda em Bacharelado em Sistemas de Informação do Centro Universitário Newton Paiva

3 NoSQL (Not only SQL) surgiram da necessidade de escalar bancos de dados para o resolver os problemas das aplicações web que operam em larga escala.

4 Secure Shell (SSh) é um protocolo de rede. Secure Copy (SCP) transfere arquivos entre hosts locais ou remotos.