Márcio Fernandes Justino

Fragmentação de Arquivos em File Carving em Sistemas de Arquivos NTFS

São Paulo – SP

Maio / 2013

Márcio Fernandes Justino

Fragmentação de Arquivos em File Carving em Sistemas de Arquivos NTFS

Projeto de pesquisa apresentado como requisito parcial para a aprovação na disciplina Metodologia do Trabalho Científico do Curso de Computação Forense da Universidade Presbiteriana Mackenzie.

Orientadora:
Ivete Irene dos Santos

UNIVERSIDADE PRESBITERIANA MACKENZIE
INSTITUTO DE COMPUTAÇÃO
PÓS GRADUAÇÃO EM COMPUTAÇÃO FORENSE

São Paulo – SP

Maio / 2013

...

Prof. xxx
Departamento xxx
Orientador

Prof. yyy
Departamento yyy

Prof. zzz
Departamento zzz

Dedico a meus pais, cujo exemplo
de honestidade e trabalho tem marcado
minha vida, à minha esposa que me apoiou
nesta caminhada e à minha filha que
acompanhou todo este trabalho
ainda no ventre da mãe.

Resumo

A proposta desta pesquisa é explanar e determinar uma melhor forma de localização de fragmentos de arquivos não alocados durante o processo de file carving em uma perícia forense digital abordando o conceito da metodologia de identificação de fragmentos de arquivos e os benefícios que o mesmo proporciona para a análise em uma investigação de uma imagem digital quando possível localizar suas partes, permitindo assim sua identificação.

Palavras-chave: NTFS, fragmentação, carving, identificação, partes, arquivos.

Abstract

...

Sumário

Lista de Figuras

Lista de Tabelas

1 Introdução

Juntamente com o avanço da tecnologia computacional e da internet veio o aumento do número de pessoas conectadas trocando informações, seja em nível pessoal ou organizacional. Segundo o Centro de Estudos sobre as Tecnologias da Informação e da Comunicação (cetic.br), o número de usuários domésticos e no trabalho tem aumentado juntamente com o tempo em que os mesmos permanecem conectados à internet. A figura 1.1 mostra a evolução desses números até o presente momento.

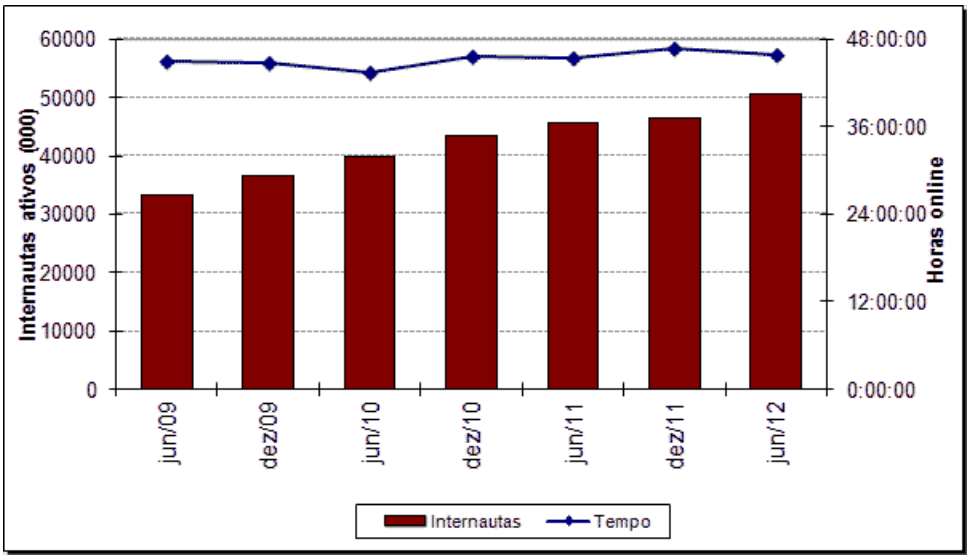

Figura 1.1: INTERNAUTAS ATIVOS EM RESIDÊNCIAS E NO TRABALHO E HORAS NAVEGADAS - 2012 (IBOPE//NETRATINGS, 2012)

Nesse meio, existem usuários que promovem o cibercrime[1] ou atividades ilegais na rede[2]. As informações computacionais são armazenadas em discos rígidos[3] usando apropriados sistemas de arquivos que são suportados pelo sistema operacional instalado no computador. Existem diversos sistemas de arquivos para armazenamento de arquivos no mercado, e um dos mais comuns atualmente é o NTFS (MAHANT, 2012).

[1] crimes cibernéticos tendo sistemas informatizados como meio de ação (CYBERCITIZEN, 2012).

[2] o termo rede será usado ao longo deste texto podendo representar a internet como um todo ou a ligação de mais de computadores entre si.

[3] unidade física de armazenamento de dados em um computador

Muitos criminosos se utilizam do artifício de excluir ou remover rastros de seus atos criminosos apagando os arquivos criados, manipulados ou alterados, acreditando que com isso seu crime seria um crime perfeito, sem possibilidade de identificação e assim da comprovação de seus atos. A técnica de file carving possibilita a análise de tais arquivos, provendo um avanço investigativo com a possiblidade de extrair provas de arquivos não mais alocados porém que ainda estejam presentes fisicamente nos discos.

A técnica de file carving é frequentemente utilizada durante investigações digitais. Conforme Memon (2011, p. S2) essa é uma técnica em que arquivos de dados são extraídos de um dispositivo digital sem o auxílio de tabelas de arquivo ou outros meta-dados do disco.

"In forensic practice, file carving can recover files that have been deleted and have had their directory entries reallocated to other files, but for which the data sectors themselves have not yet been overwritten (GARFINKEL, 2007)."[4]

1.1 Justificativa

Segundo Memon (2011, p. S2), um dos primeiros desafios em file carving pode ser encontrado na tentativa de se recuperar arquivos fragmentados. O processo de file carving é de suma importância para a investigação forense computacional e envolve a identificação de arquivos perdidos, corrompidos ou removidos do equipamento investigado. A dispersão desses arquivos não mais indexados pela tabela de alocação de arquivos do sistema de arquivos NTFS torna o processo de identificação dos arquivos um desafio para a investigação e identificação de ilícitos.

> During a digital forensic investigation many diferent pieces of data are preserved for investigation, of which bit-copy images of hard drives are the most common. These images contain the data allocated to files as well as the unallocated data. The unallocated data may still contain information that is relevant to an investigation, in the form of (parts of) intentionally deleted or automatically removed temporary files. Unfortunately, this data is not always easily accessible: a string search on the raw data might recover (parts of) interesting text documents, but it won't help to get to information present in for example images or compressed files. Besides that, the exact strings to look for may not be known beforehand. To get to this information, the deleted files have to be recovered (KLOET, 2007).[5]

[4]Na prática forense, file carving pode recuperar arquivos que tenham sido apagados e tenham suas entradadas de diretório realocadas para outros arquivos, desde que seus setores de dados ainda não tenham sido sobrescritos.

[5]Durante uma investigação forense digital, muitas peças diferentes de dados são preservadas para investigação, das quais imagens de discos rígidos (HD's) são as mais comuns. Essas imagens contêm os dados alocados para arquivos, bem como os dados não alocados. Os dados não alocados ainda podem conter informações relevantes para uma investigação, sob a forma de (partes de) intencionalmente excluídos ou arquivos temporários removidos automaticamente. Infelizmente, esses dados nem sempre são facilmente acessíveis: uma sequência de caracteres da

1.2 Fragmentação de Arquivos

Segundo Nasir Menom (2011, p. S2), um dos primeiros desafios do processo de investigação utilizando file carving é justamente a tentativa de recuperar os framentos de arquivos não alocados. A fragmentação de arquivos é um desafio para o processo de file carving e é de suma importância para a recuperação de arquivos perdidos em processos de investigação digital. Tendo em vista o presente desafio tem-se a necessidade de se determinar qual a melhor técnica para identificação de fragmentos de arquivos no processo de file carving.

1.3 Hipótese(s)

Análise dos metadados de um arquivo quanto ao seu início e fim (cabeçalho e rodapé do arquivo), informações que determinam onde os dados de um arquivo começam e onde terminam. Identificar um padrão de dados na localização dos fragmentados de forma consistente, reduzindo assim os falsos positivos comumente apresentados no processo de file carving permitirão a identificação máxima de conteúdo do arquivo no processo de investigação forense.

1.4 Objetivo Geral

Determinar uma melhor metodologia de localização de fragmentos de arquivos no processo de file carving em sistemas de arquivos NTFS.

1.5 Objetivo Específico

Para chegar ao objetivo principal e determinar uma melhor metodologia de localização de fragmentos de arquivos é necessário entender primeiramente e de forma mais detalhada alguns items específicos:

- Verificar como são identificados os arquivos no sistema de arquivo NTFS;
- Verificar como um arquivo fragmentado é armazenado em um sistema NTFS;
- Levantar uma padronização entre os fragmentos de arquivos para melhor localização;

pesquisa sobre os dados brutos pode recuperar (partes de) documentos de texto interessantes, mas ele não vai ajudar para obter a informação presente em, por exemplo, imagens ou arquivos compactados. Além disso, as sequências de caracteres exatas para procurar não podem ser conhecidas antecipadamente. Para obter esta informação, os arquivos apagados precisam ser recuperados.

- Identificar formas de localização de fragmentos dos arquivos não alocados;

Verificar assim a forma como os arquivos são registrados nos sistemas de arquivos NTFS, o processo de diferenciação de tipos de arquivos para determinar o início e o fim de um arquivo (área de cabeçalho, área de dados, de metadados e ponto de fim de arquivo), podendo então encontrar certos padrões que possam permitir a identificação de partes de um arquivo fragmentado no sistema de arquivos.

1.6 Metodologia

...em desenvolvimento...

2 *Levantamento Bibliográfico*

Desenvolvimento do texto baseado nas referências pesquisadas, que corresponde ao corpo teórico do trabalho. Aqui, deve-se incluir a pesquisa bibliográfica dos assuntos relacionados ao tema, obtendo o corpo de conhecimentos que balizará o estudo. Os capítulos e subtítulos devem ser encadeados de forma a explicar, discutir e demonstrar o conteúdo do trabalho. Esta etapa consiste em estruturar de maneira lógica as partes do trabalho, seus capítulos, de forma que estes estejam bem inseridos no contexto do discurso e da redação. Pode ser que, no decorrer do trabalho, esta construção lógica precise sofrer modificações ou desdobramentos, até que se chegue ao plano definitivo (Severino: 1979, p. 86). Toda a argumentação e raciocínio são construídos em cima das leituras, experiências e da vivência intelectual a respeito do problema da pesquisa. Esta fase pressupõe o levantamento de toda a documentação existente sobre o assunto da pesquisa: livros, artigos, revistas. Conforme colocado por Severino (1979, p. 81), esta é a ?fase da heurística, ciência, técnica e arte da pesquisa de documentos?. É importante salientar que o trabalho, nesta etapa, pressupõe uma finalidade didática, não podendo ser uma pura criação mental do aluno. Deve se balizar em pesquisas e consultas de documentação, em livros, artigos técnicos e jornais, de forma a gerar quantidade suficiente de informação sobre o tema do trabalho. As idéias e opiniões dos alunos, da fase do levantamento bibliográfico, podem ajudar na ligação entre idéias de autores, mas não devem prevalecer sobre a pesquisa dos autores. É interessante que, antes de iniciar a leitura dos materiais, o aluno elabore um roteiro de trabalho (que podem ser a origem dos capítulos), ou seja, uma primeira estruturação, de forma que a leitura e a pesquisa se dêem dentro deste roteiro. Ou seja, tenha em mente quais são as colunas mestras do trabalho que demonstrarão, dos textos lidos, os elementos que devem ser retidos para o aproveitamento na composição do trabalho. Vale salientar que este roteiro é provisório, podendo ser reformulado no decorrer do trabalho. O aluno deve proceder uma análise de possíveis materiais que podem ser interessantes para o estudo do tema e consultar seu orientador sobre livros que são imprescindíveis para a consecução do estudo. Algumas dicas válidas: tire um tempo para ir a biblioteca e analise os livros pelo seu índice. Muitas vezes, importantes teorias não aparecem no título da obra, mas sim em capítulos dos livros. Pesqui-

sar com paciência, nesta fase, é a chave do sucesso futuro do trabalho. Procure analisar livros relacionados com o assunto, revistas, teses. Com certeza, esta atitude vai ajudar não apenas no estudo e detalhamento do tema em si, como também dar um foco mais interessante ao estudo ou ao problema de pesquisa. Lembre-se: pesquisar não significa perder tempo, mas sim, ganhar luz e maior proficiência sobre o assunto. Uma última colocação acerca do levantamento bibliográfico: atualmente, os alunos tendem a pesquisar com ênfase errônea na Internet. Apesar de este ser um meio de informação altamente eficaz, certamente não é suficiente para gerar o conhecimento necessário ao desenvolvimento teórico do trabalho. Podemos comparar a Internet a um lago, que pode ser extenso, porém é raso, superficial. Outro cuidado diz respeito à ação ?copiar e colar?. Nesta hora, o aluno deve ter amadurecimento necessário para não copiar trechos de obras, mesmo que contenha a citação da mesma, pois isso se configura plágio. O trabalho deve ser o de ler e reformular o que foi dito pelo autor, dentro da perspectiva de seu tema. Um exemplo de reformulação: ?Segundo Coda (ano), a motivação pode ser definida como...?. Perceba que, desta forma, o trecho do autor Coda não esta sendo copiado, mas comentado pelo autor da monografia. Quando for absolutamente imprescindível copiar definições ou trechos de obras, este deve ser colocado entre aspas, de forma que fique clara a autoria de determinado trecho, e deve ser indicada a página do livro de onde foi retirada. Trechos com mais de 4 linhas devem vir em fonte menor e com margem esquerda de 4 cm. A seguir, devem ser desenvolvidos os capítulos. Um ponto importante é fazer a ligação entre um capítulo e o seguinte, de forma que o leitor entenda completamente a sequenciação.

3 Sistemas de Arquivos NTFS

...

3.1 Conceito

...

3.2 Estrutura de Arquivos

...

3.3 Alocação de Arquivos

...

4 File Carving

...

4.1 Conceito

...

4.2 Assinatura de Arquivo

...

4.3 Número Mágico

...

5 File Carving Avançado

...

5.1 Conceito

...

5.2 Funcionamento

...

5.3 Fragmentação

...

5.4 Ponto de Fragmentação

...aqui serão incluídos também algoritmos de detecção de pontos de fragmentação como listado no mind map (Sequential Hypothesis Testing)...

6 Considerações Finais

...em desenvolvimento...

Referências Bibliográficas

CYBERCITIZEN. *What is Cyber Crime?* cybercitizenship, 2012. Disponível em: <http://www.cybercitizenship.org/crime/crime.html>.

GARFINKEL, S. L. Carving contiguous and fragmented files with fast object validation. *ELSEVIER*, v. 4S, p. S2–S12, 2007.

IBOPE//NETRATINGS, N. *Painel IBOPE/NetRatings*. cetic.br, 2012. Disponível em: <http://www.cetic.br/usuarios/ibope/w-tab02-01-cons.htm>.

KLOET, S. *Measuring and Improving the Quality of File Carving Methods*. Dissertação (Mestrado) — Eindhoven University of Technology, 10 2007.

MAHANT, B. S. H. Ntfs deleted files recovery - forensics view. *IRACST - International Journal of Computer Science and Information Technology and Security (IJCSITS)*, v. 2, n. 3, p. 491–497, 2012.

Anexos

...