Análise de Ferramentas Forenses de Recuperação de Dados

Josilene dos Santos Nascimento, Klarissa de Souza Jerônimo e Pedro Chaves de Souza Segundo

***Abstract*—This paper is an analysis of forensic data recovery tools and it aims to establish how and in which situations such tools may be used. Post Mortem was the type of forensic analysis chosen in this work, and disk images were used for testing. Three tools were analyzed: Foremost, Autopsy/The Sleuth Kit and FTK Imager, one of them proprietary and the others, open source.**

***Index Terms*—Data recovery, forensic data recovery tools, open source.**

I. Introdução

Numa cena de crime, um bom lugar para procurar vestígios de atividades ilícitas e para traçar um perfil da pessoa investigada são os equipamentos digitais de que ela faz uso. No entanto, não apenas os dados que estão aparentes são importantes. Muitos dados que podem vir a ser vitais para um caso investigado podem ter sido apagados propositadamente na intenção de evitar problemas futuros ou por não serem mais de interesse do proprietário. Acrescente-se a isso o fato de que muitos dos atacantes de computadores, fazendo uso de técnicas anti-forenses, apagam os arquivos utilizados no intuito de esconder o registro de suas atividades.

Segundo Vacca [1], os arquivos apagados de um sistema não são necessariamente sobrescritos e ainda podem ser recuperados usando ferramentas e procedimentos adequados.

Este trabalho tem como objetivo principal fazer uma análise de ferramentas forenses quanto à capacidade de recuperação de arquivos apagados. As ferramentas escolhidas são Foremost e Autopsy/The Sleuth Kit, disponíveis em software livre. Adicionalmente, incluiu-se o FTK Imager, ferramenta proprietária utilizada em perícias forenses de informática, distribuído gratuitamente pela empresa AccessData.

A abordagem escolhida para os testes foi a de comparar o desempenho de cada ferramenta frente a diferentes sistemas de arquivos (Ext3, Ext4, Fat32, Ntfs) nas seguintes situações ou cenários:

1. Arquivos simplesmente apagados;
2. Arquivos apagados e disco parcialmente sobrescrito;
3. Arquivos apagados e disco totalmente sobrescrito.

Os arquivos utilizados foram figuras de extensões jpg, bmp, gif e png, no total de 800 arquivos, cujos códigos de integridades (*hashes*) foram calculados para comparação com os resultados. Essas figuras foram gravadas em partições, onde foram simulados os três cenários. As imagens (retratos das partições) resultantes desses cenários são alvo das ferramentas forenses. De acordo com os *hashes* obtidos dos arquivos recuperados, há dois tipos de recuperação: total, quando há coincidência com o *hash* inicial, e parcial, quando o arquivo recuperado não tem *hash* presente na lista dos *hashes* originais.

Os experimentos com as ferramentas Foremost e TSK/Autopsy foram realizados sobre uma plataforma Linux Ubuntu 9.04, Kernel 2.6.28-17, e para os testes com a ferramenta FTK Imager, foi utilizado Windows XP, Service Pack 3. Ambos os sistemas usam uma plataforma de 32 bits.

Exames qualitativos [2] mostraram que as ferramentas têm capacidade de recuperação dependente do sistema de arquivos. Neste artigo, será observada a capacidade de recuperação de figuras pelas ferramentas em alguns cenários.

Os resultados dos experimentos realizados sugerem ferramentas forenses mais adequadas para cada tipo de sistema de arquivos e cenário, auxiliando os peritos na escolha de suas ferramentas dependendo dos casos em que estiverem envolvidos.

II. Exames E Metodologia

A. Montando os Sistemas de Arquivos

Os sistemas de arquivos escolhidos para os testes foram: Ext3, Ext4, FAT32 e NTFS.

A instalação dos sistemas de arquivos para os testes foi feita criando-os em arquivos dentro do sistema em uso e montando-os como partições. Estas partições foram tomadas como base para o preparo das imagens que foram submetidas às ferramentas de análise forense. O passo a passo para a instalação e montagem de cada um deles é descrito na Tabela 1, seguindo um modelo de Jones [3].

Foram criadas partições de 600MB de forma a facilitar a manipulação dos arquivos dentro delas e a posterior análise dos dados.

Os comandos da Tabela 1 realizam para cada sistema de arquivos e tamanho da imagem:

1. Wipe da partição, escrevendo zero em todo o seu conteúdo;
2. Criação do sistema de arquivos na partição, utilizando todo o seu tamanho;

3. Montagem da partição no sistema de arquivos, tornando-a disponível para uso no sistema operacional, onde será possível copiar e apagar arquivos.

TABELA I
PASSO A PASSO DA CRIAÇÃO E MONTAGEM DAS PARTIÇÕES UTILIZADAS NOS TESTES

Sistema de Arquivos	Comandos
Ext3	dd if=/dev/zero of=ext3.img bs=1k count=614400 losetup /dev/loop0 ext3.img mkfs.ext3 -c /dev/loop0 mount -t ext3 /dev/loop0 /mnt/disk-1/
Ext4	dd if=/dev/zero of=ext4.img bs=1k count=614400 losetup /dev/loop1 ext4.img mkfs.ext4 -c /dev/loop1 mount -t ext4 /dev/loop1 /mnt/disk-2/
Fat32	dd if=/dev/zero of=fat32.img bs=1k count=614400 losetup /dev/loop2 fat32.img mkfs.vfat -F 32 /dev/loop2 mount -t vfat /dev/loop2 /mnt/disk-3/
NTFS	dd if=/dev/zero of=ntfs.img bs=1k count=614400 losetup /dev/loop3 ntfs.img mkfs.ntfs /dev/loop3 mount -t ntfs /dev/loop3 /mnt/disk-4/

Nessas partições montadas foram copiados 800 figuras, sendo 200 de cada tipo: jpeg, png, gif e bmp. Esse conjunto soma ao todo 278 MB. Por questão de uniformidade, o mesmo conjunto de arquivos é usado para todos os testes. Esses arquivos foram copiados para as partições de forma a ficarem espalhados homogeneamente quanto à extensão, seguindo a sequência jpg, png, bmp e gif exemplificada a seguir: fig0001.jpg, fig0001.png, fig0001.bmp, fig0001.gif, fig0002.jpg, fig0002.png, fig0002.bmp, fig0002.gif, [...] fig0199.jpg, fig0199.png, fig0199.bmp, fig0199.gif, fig0200.jpg, fig0200.png, fig0200.bmp, fig0200.gif.

B. Cenários

Como o objetivo é a recuperação de arquivos apagados do disco, foram preparados três cenários diferentes para a realização dos testes, cada um simulando uma situação em que os arquivos a serem recuperados poderiam ser expostos dentro de um sistema de arquivos.

Cenário 1

Este cenário exemplifica o caso em que um arquivo tenha sido apagado do disco e a sua recuperação tenha sido iniciada sem que o sistema tenha sofrido alterações. O objetivo é testar o desempenho das ferramentas e o comportamento dos sistemas na tentativa de recuperar os arquivos recentemente apagados.

Foram usadas as partições formatadas de 600MB e em cada uma delas feito o seguinte processo:

- Cópia dos 800 arquivos;
- Desmontagem/montagem;
- Remoção dos arquivos;
- Desmontagem/montagem.

Após a remoção dos arquivos, o processo de desmontagem/montagem é repetido para apagar qualquer resquício que possa ter sobrado no *buffer* dos sistemas. Isto poderia comprometer os resultados dos testes caso a ferramenta recuperasse um arquivo não por ter podido retirá-lo de dentro da partição, mas por ele ainda estar na memória do sistema.

Por fim, obtêm-se as imagens para os exames, executando o comando a seguir para cada um dos sistemas de arquivos:

dd if=/dev/loop0 of=<sistema>_<cenário>.dd

Cenário 2

Neste cenário, para sobrescrever a partição, são utilizados arquivos diversos que cheguem a um tamanho por volta dos 300MB, que é um pouco mais do que o tamanho do total dos arquivos a serem recuperados. Como as partições a serem usadas são de 600MB, sobra espaço suficiente para acomodar estes novos arquivos sem que os demais sejam tocados. Ou seja, dependendo da formatação e da manipulação dos dados do sistema testado, os arquivos que deveriam ser recuperados podem ter sido sobrescritos ou não, respectivamente dificultando e facilitando o trabalho das ferramentas de recuperação de dados aqui testadas.

Neste Cenário 2, foi repetido o processo de criação do Cenário 1 e incluído após a remoção dos arquivos o seguinte passo:

- Cópia de novos arquivos chegando a um total de 300MB.

Neste cenário, as partições de onde os arquivos foram apagados no Cenário 1 são agora preenchidas com novos arquivos, sem relação com os primeiros, de forma tal que a partição pareça 100% ocupada.

Neste caso, os arquivos escritos e apagados no Cenário 1 têm uma chance quase total de terem sido sobrescritos. Busca-se, então, verificar se as ferramentas conseguem ainda recuperar algum arquivo ou fragmento.

Para este terceiro cenário, às partições antes usadas para compor o Cenário 1, após a remoção dos arquivos, foi incluído o seguinte passo:

- Cópia de novos arquivos de forma a preencher todo o espaço da partição.

C. Descrição das Ferramentas

Será apresentado nesta seção como foram utilizadas as ferramentas de recuperação para realização dos exames.

Ferramenta 1 - Foremost

Os exames com o Foremost foram realizados através de uma linha de comando para cada imagem. Nenhuma configuração adicional foi feita na ferramenta depois de sua instalação, que também seguiu o padrão.

Para executar a ferramenta e iniciar o teste, foi usado o comando a seguir, o qual foi repetido para cada sistema e cenário:

foremost -i <imagem> -o <pasta_de_destino>

Com este comando, todos os tipos de arquivos suportados pela ferramenta seriam buscados dentro da imagem especificada.

O resultado desta ferramenta para cada cenário e sistema é uma pasta contendo os arquivos recuperados, separados em subpastas de acordo com a extensão.

Ferramenta 2 - TSK/Autopsy

Os exames realizados com o TSK são feitos através de uma ferramenta gráfica via web, o Autopsy, ativada através do comando ./autopsy executado na pasta onde a ferramenta foi compilada. Para iniciar os testes, adicionam-se as imagens a serem analisadas, utilizando os menus da ferramenta.

Para obtenção dos arquivos possivelmente recuperados, foi utilizada a análise "File Type". Neste caso, clica-se em "Sort Files by Type", fazem-se as escolhas necessárias na tela que se segue e, depois de confirmadas as opções, abre-se uma tela onde é mostrado um resumo do que foi recuperado. Se a opção de gravar os arquivos em disco foi marcada, todos os arquivos recuperados podem ser conferidos dentro da pasta de análise do Autopsy no diretório explicitado na página de resumo. As imagens podem ser visualizadas pelo navegador mesmo sem terem sido salvas em disco ao seguir o link indicado nesta página de análise do Autopsy.

Ferramenta 3 – FTK Imager

A ferramenta é ativada simplesmente com um clique duplo sobre o seu arquivo executável, "FTK Imager.exe". Dentro de sua interface gráfica, para incluir uma nova imagem a ser analisada, basta ir ao botão "Add Evidence Item" e na janela que se abre, escolher o item "Image File", clicar em Avançar, escolher a imagem e clicar em concluir.

A imagem é então incluída pela ferramenta, que tenta recuperar todos os dados da partição ou disco do qual a imagem foi feita.

III. Resultados

Os resultados foram separados em três tipos:

- Total: quando o *hash* é idêntico ao do arquivo original. Neste caso, o seu conteúdo pode ser visto por completo. O resultado foi encaixado nesta categoria mesmo quando as ferramentas não conseguiram recuperar os metadados dos arquivos, como nomes, datas.
- Parcial: quando o *hash* calculado não coincide com nenhum dos originais. Neste caso, o arquivo recuperado pode ser visualizado em pedaços. Exemplo: parte de uma figura.
- Nulo: quando nada do que é recuperado for válido, ou seja, quando não é possível ser visualizado nenhum fragmento ou metadado dos arquivos.

Seguem observações gerais sobre o comportamento das ferramentas utilizadas:

- O FTK Imager não reconheceu o sistema Ext4;
- Algumas ferramentas não conseguem recuperar metadados em alguns sistemas e cenários.
- O Foremost não consegue recuperar metadados para nenhum sistema de arquivos.

Os resultados obtidos são apresentados em gráficos por cenário e sistema de arquivos. Nesses gráficos, o número de arquivos recuperados pelas ferramentas é indicado na vertical e as ferramentas, na horizontal. As barras representam o tipo de recuperação, conforme informado nesta seção.

Cenário 1

O desempenho das ferramentas ao tentar a recuperação dos arquivos apagados de partições formatadas com os sistemas Ext3, Ext4, FAT32, e NTFS é mostrado, respectivamente, nas figuras 1 a 4.

Para o sistema Ext3, mostrado na Fig. 1, a ferramenta Foremost foi a que obteve o melhor desempenho, conseguindo recuperar pouco mais da metade dos arquivos. No entanto, deve-se observar que os arquivos foram recuperados sem seus nomes originais. O Autopsy conseguiu a recuperação apenas dos nomes dos arquivos, enquanto o FTK Imager não chegou a recuperar nada.

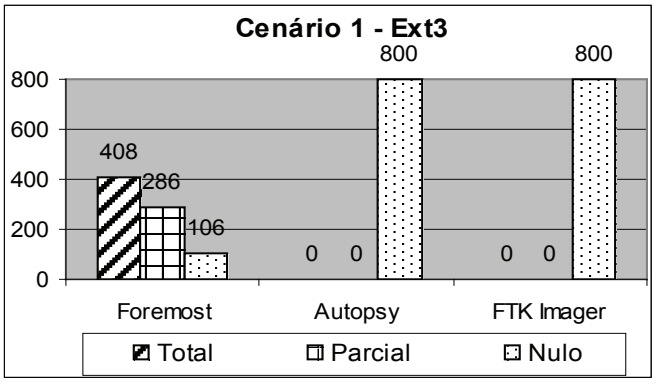

Fig. 1. Quantidade de arquivos recuperados para o Cenário 1, usando o sistema Ext3

Os resultados para o sistema de arquivos Ext4, mostrados na Fig. 2, mostram uma sensível melhora na recuperação da ferramenta Foremost, onde esta chega a recuperar quase 90% dos arquivos. As demais ferramentas mostraram um resultado semelhante com o obtido para o sistema Ext3.

No caso de também ser necessária a recuperação dos nomes dos arquivos, sugere-se usar os resultados do Autopsy como um complemento aos obtidos pelo Foremost.

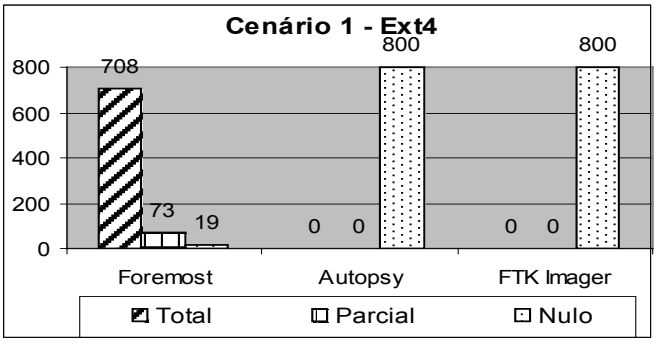

Fig. 2. Quantidade de arquivos recuperados para o Cenário 1, usando o sistema Ext4

O desempenho das ferramentas sofreu uma melhora significativa, como demonstrado na Fig. 3, ao analisar uma imagem do sistema FAT32. Para esta imagem, a ferramenta FTK Imager teve 100% de aproveitamento, a Foremost obteve uma recuperação quase total e o Autopsy recuperou cerca de metade dos arquivos.

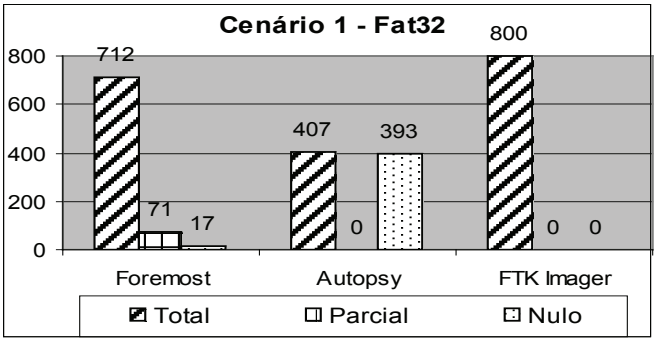

Fig. 3 . Quantidade de arquivos recuperados para o Cenário 1, usando o sistema FAT32

O resultado ilustrado na Fig. 4 para o sistema NTFS neste cenário mostra resultados semelhantes ao anterior para as ferramentas Autopsy e Foremost. Entretanto, para o FTK Imager o resultado foi completamente diferente, com recuperação de apenas poucos arquivos (~3%) e ainda assim parcialmente.

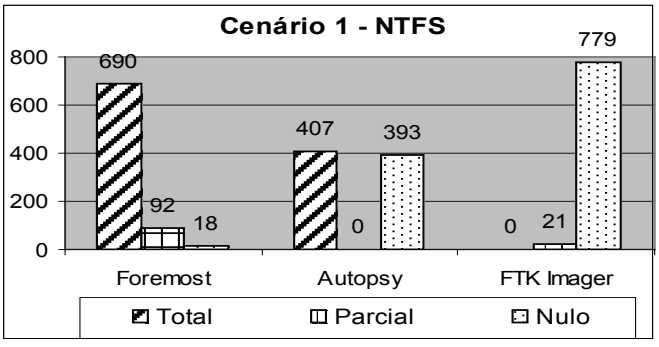

Fig. 4 . Quantidade de arquivos recuperados para o Cenário 1, usando o sistema NTFS

Nesse cenário, verificou-se que apesar dos arquivos serem apenas apagados, não se pode sempre recuperar todos os arquivos. Em geral, a ferramenta Foremost se mostrou superior às demais para o Cenário 1. Como pode ser observado, as ferramentas tendem a recuperar melhor os arquivos quando gravados em sistemas FAT32, onde o FTK Imager foi o melhor com recuperação total de 100% dos arquivos.

Cenário 2

Neste cenário, foram usadas as mesmas imagens do Cenário 1, onde após os arquivos serem apagados, arquivos que completam um total de 300MB foram copiados de forma a sobrescrever a partição. A intenção foi deixar espaço suficiente para que, dependendo do modo de operar do sistema de arquivos usado, os arquivos a serem recuperados não sejam todos sobrescritos. Aqui também pode ser observado o comportamento dos sistemas de arquivos.

O desempenho das ferramentas frente a uma imagem do sistema Ext3 submetida ao Cenário 2 é mostrado na Fig. 5. Comparando os resultados com os do Cenário 1, foram identificadas mudanças apenas para a ferramenta Foremost. Esta recuperou menos da metade dos arquivos que no cenário anterior. As ferramentas Autopsy e FTK Imager mantiveram uma recuperação de dados nula.

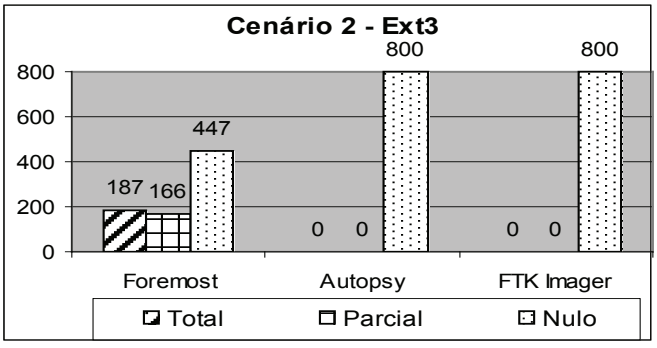

Fig. 5. Quantidade de arquivos recuperados para o Cenário 2, usando o sistema Ext3

Para o sistema Ext4, ilustrado no Fig. 6, o desempenho da Foremost cai sensivelmente, menos de 6% dos arquivos são recuperados totalmente. As ferramentas FTK Imager e Autopsy continuam não recuperando arquivos.

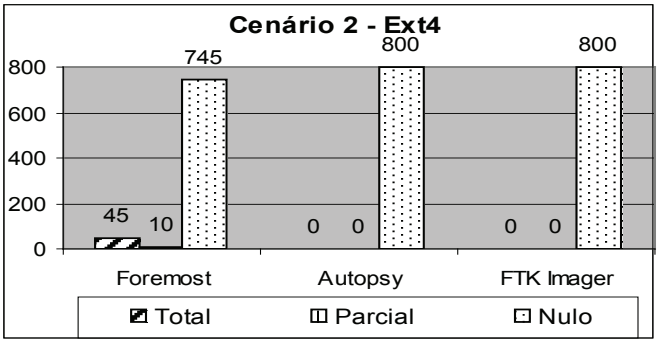

Fig. 6. Quantidade de arquivos recuperados para o Cenário 2, usando o sistema Ext4

O desempenho na recuperação de arquivos para o sistema FAT32 neste cenário é muito semelhante ao que foi observado para o cenário anterior, como pode ser visto na Fig. 7. As ferramentas Foremost e FTK Imager recuperaram quase a totalidade dos dados, enquanto o Autopsy recuperou cerca da metade.

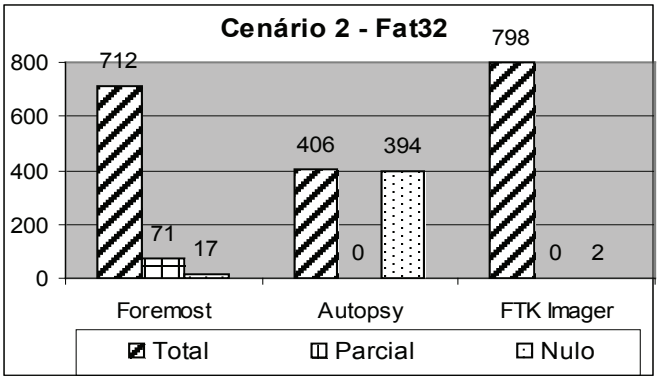

Fig. 7. Quantidade de arquivos recuperados para o Cenário 2, usando o sistema FAT32

Com relação ao sistema NTFS, mostrado na Fig. 8, todas as ferramentas têm o seu desempenho reduzido. Cerca de 35% dos arquivos são recuperados com a ferramenta que obteve os melhores resultados, o Foremost. O Autopsy também apresentou uma recuperação regular (21%) e o FTK Imager recuperou poucos arquivos e só parcialmente.

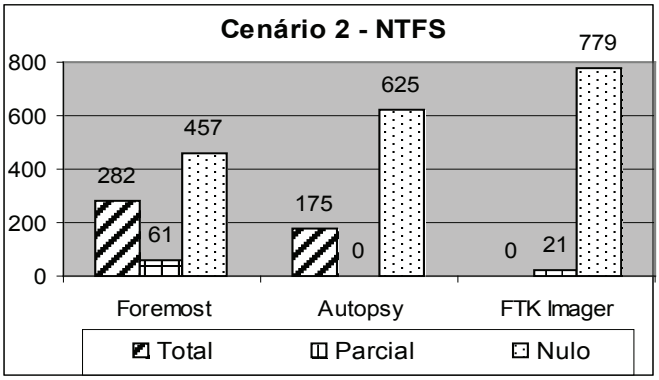

Fig. 8. Quantidade de arquivos recuperados para o Cenário 2, usando o sistema NTFS

Ao comparar os resultados deste cenário com o anterior, pode-se observar que houve um decréscimo no desempenho das ferramentas, com exceção para a recuperação do sistema FAT32, que permaneceu o mesmo.

Observa-se também que a ferramenta Foremost ainda é a que obtém os melhores resultados, exceto no sistema de arquivos FAT32, onde o FTK Imager é melhor.

Cenário 3

Neste cenário, foram copiados novos arquivos para as partições já usadas antes para os testes do Cenário 1 de forma que ocupassem todo o espaço disponível. Espera-se medir a capacidade de recuperação das ferramentas quando os arquivos a serem recuperados tenham sido sobrescritos por outros dados.

Como pode ser visto nas figuras desta seção, numeradas de 9 a 12, o desempenho das ferramentas foi quase sempre nulo. A ferramentas FTK Imager apresentou um desempenho nulo para todos os sistemas testados. O Autopsy e o Foremost demonstraram um desempenho semelhante, conseguindo recuperar apenas alguns arquivos no sistema NTFS, inclusive com recuperação total de arquivos.

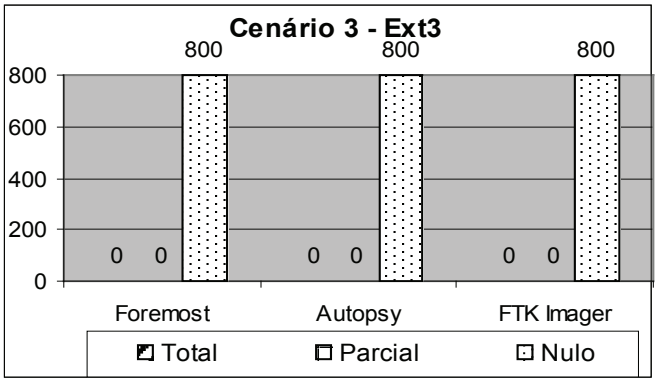

Fig. 9. Quantidade de arquivos recuperados para o Cenário 3, usando o sistema Ext3

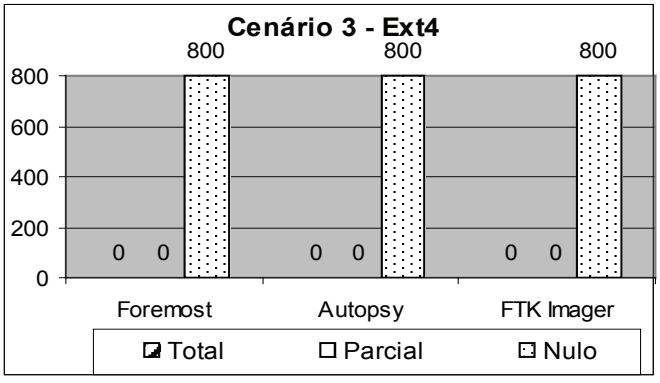

Fig. 10. Quantidade de arquivos recuperados para o Cenário 3, usando o sistema Ext4

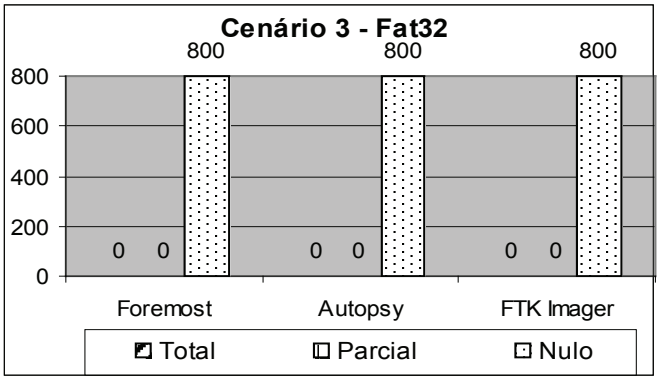

Fig. 11. Quantidade de arquivos recuperados para o Cenário 3, usando o sistema FAT32

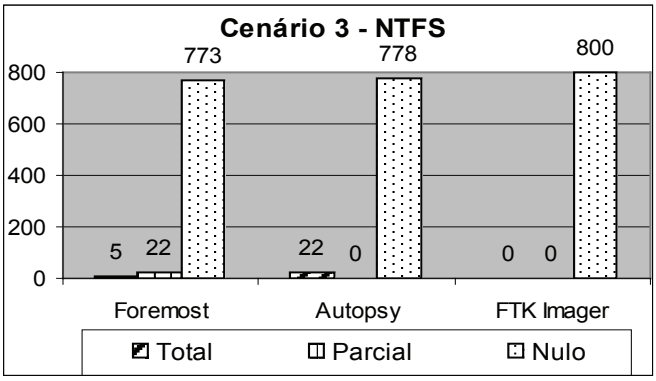

Fig. 12. Quantidade de arquivos recuperados para o Cenário 3, usando o sistema NTFS

Do Cenário 3, de onde os arquivos foram apagados e sobrescritos com dados que preencheram todo o tamanho da partição, esperava-se que o conteúdo de todos os arquivos fosse sobrescrito, provocando resultados nulos exclusivamente. Esse comportamento foi observado para todos os sistemas de arquivos, com exceção do NTFS, onde se obteve a recuperação parcial de 2,75% dos arquivos na ferramenta Foremost, e ainda recuperação total de 0,62% (5 arquivos) e 2,75% (22 arquivos) para o Foremost e o Autopsy, respectivamente.

A recuperação parcial para este cenário também foi observada na referência [2] com recuperação de metadados dos arquivos originais, sendo explicada pelos metadados dos novos arquivos gravados não ocuparem todo o espaço na tabela de partição antes preenchido pelos metadados dos arquivos apagados. No entanto, neste trabalho a recuperação parcial refere-se a fragmentos dos arquivos e não somente a metadados.

Os autores não identificaram justificativa para a recuperação parcial e total no cenário 3.

IV. Conclusão

Ao submeter as ferramentas forenses Foremost, Autopsy e FTK Imager a três cenários diferentes de perda de dados, pode-se concluir que é relativamente simples recuperar dados que tenham sido apagados imediatamente do disco. No entanto, à medida que o sistema é usado e novos dados são inseridos nele, fica cada vez mais difícil conseguir esta recuperação.

Como apresentado neste trabalho, nenhuma dessas ferramentas é totalmente eficiente para todos os sistemas de arquivos ou cenários em que os seus dados possam estar inseridos. A melhor abordagem é, antes de iniciar o procedimento de recuperação, observar qual o sistema de arquivos em que a partição analisada é formatada.

O que pode ser observado pelos resultados obtidos neste trabalho, com amostra de arquivos de figuras, é que a ferramenta Foremost foi a que obteve um melhor desempenho no geral se comparada às demais, excetuando-se apenas no caso do sistema FAT32, onde a ferramenta FTK Imager foi melhor.

Deve ser levado em consideração o fato de que a ferramenta Foremost não recupera o nome dos arquivos. A solução proposta é que seja usada em conjunto com o Autopsy, pois quase sempre são recuperados os metadados completos dos arquivos, apesar dos dados estarem perdidos.

Referências

[1] VACCA, John R. Computer Forensics: Computer Crime Scene Investigation. 2a ed. Boston: Charles River Media, 2005.

[2] Nascimento, Josilene dos Santos. Análise de Ferramentas Forenses de Recuperação de Dados. João Pessoa, 2010. Monografia (Curso de Segurança da Informação) – Faculdade de Tecnologia de João Pessoa – FATEC.

[3] JONES, M. Tim. Anatomia do Sistema de Arquivos do Linux. Disponível em: <http://www.ibm.com/developerworks/br/library/l-linux-filesystem/index.html>. Acesso em: 30 jan. 2010.