Breve abordagem aos fundamentos e metodologias da Informática forense

Introdução

É hoje incontestável que a maior parte da informação produzida, é criada e armazenada no formato digital e estima-se que mais de metade da documentação relacionada com a actividade económica, nunca deixará o domínio digital.

Isto significa que os documentos em formato papel, associados ao mundo dos negócios, constituem apenas uma pequena parte, sendo já significativamente maioritário, o número de documentos em formato digital.

Esta realidade contrasta com o domínio que o papel continua a exercer no campo da justiça, onde predomina ainda este formato.

Não sou eu que o digo.... a esta conclusão chegaram há muito dois ilustres estudiosos destas matérias, **Fred e Christine Galves,** conforme documenta o seu artigo: "*Ensuring the Admissibility of Electronic Forensic Evidence and Enhancing Its Probative Value at Trial*", publicado na revista "Criminal Justice Magazine.

Ciência Forense

As ciências forenses têm por objectivo isolar elementos que permitam reconstruir um crime ou incidente, visando identificar suspeitos, apreender culpados, ilibar inocentes, e entender as motivações que levaram à prática do crime. No âmbito da informática forense, esses elementos são de natureza digital e residem em dispositivos electrónicos de diversas naturezas.

O princípio da troca de Locard que está na base da generalidade das ciências forenses é igualmente válido na Informática Forense não apenas na envolvente física mas sobretudo na vertente digital pois a actividade desenvolvida sobre um computador, por mais simples que seja, cria artefactos que ficam gravados no seu disco rígido que constituem marcas irrefutáveis da acção desenvolvida.

Informática Forense

"Ramo das ciências forenses que, recorrendo a ferramentas e metodologias cientificamente comprovadas, assegura a identificação, preservação, recolha, validação, análise, interpretação, documentação e apresentação de evidências digitais, obtidas a partir de qualquer dispositivo que armazene ou processe informação no formato digital, com o objectivo de facilitar ou favorecer a reconstrução de eventos relacionados com práticas criminalizáveis".

O "mundo" físico que nos rodeia, é caracterizado por uma natureza determinística e finita, no qual, propriedades intangíveis tais como o tempo, o espaço, a identidade ou a localização física surgem como inalteráveis, encontrando-se qualquer delas fora do nosso controlo.

No "Mundo Digital" as acções são virtualmente independentes quer do tempo quer da localização física e no qual, com adequados conhecimentos, é possível alterar cada uma das propriedades antes referidas.

Por outro lado, no "mundo" físico, o investigador pode observar directamente muitos estados e eventos, usando os seus próprios sentidos, enquanto, no "mundo" digital, essa observação é indirecta pois só com recurso a hardware e software adequados, é possível observar (interpretar) eventos e estados digitais.

Fundamentos da Informática Forense

Num computador ou dispositivo semelhante, quer o respetivo sistema operativo quer as aplicações informáticas executadas, geram e armazenam muito mais informação do que aquela que o utilizador visualiza.

Trata-se, na maior parte dos casos de informação ativa, podendo estar acessível ao utilizador ou residir em localizações obscuras, por vezes em formatos codificados que impossibilitam a respectiva interpretação. Dispondo dos conhecimentos e ferramentas adequadas, essa mesma informação pode revelar-se esclarecedora quando interpretada e correlacionada.

Quando acessível ao utilizador, a respectiva interpretação normalmente requer conhecimento especializado, caso dos “Metadados”. (Os metadados fornecem os recursos necessários para entender os dados através do tempo. Informações de como os dados foram criados/derivados, ambiente em que residem e/ou residiram, alterações feitas, entre outras, são obtidas a partir dos metadados). Por exemplo, o Microsoft Outlook regista a data em que é criado um dado contacto, mas poucos de nós personaliza o programa para que mostre o item "data de criação", contudo este item existe e pode revelar-se de grande utilidade num processo de investigação, permitindo por exemplo confirmar com enquadramento temporal, uma ligação entre dois indivíduos.

Quanto à informação residente em locais obscuros, temos como exemplo os ficheiros de log, ficheiros de sistema ocultos e informação registada de forma codificada cuja análise pode inclusivamente permitir tirar ilações quanto ao comportamento do utilizador.

Para além da informação “Ativa” há ainda a considerar vastas regiões dos dispositivos de armazenamento que não estão acessíveis quer ao sistema operativo quer às aplicações, as quais podem constituir importantes repositórios de evidências pois constituem uma espécie de "aterros sanitários de dados", referenciados como “Unallocated clusters” e “Slack Spaces”, onde se podem encontrar muitos dos elementos que, quer utilizadores, quer sistema operativo e aplicações, foram descartando ao longo do período de vida útil do equipamento.

Aceder e interpretar este vasto conjunto de dados não estruturados, exige o recurso a ferramentas especializadas bem como técnicas e competências específicas, e constitui o foco da Informática Forense.

A Informática forense apresenta-se assim como um processo especializado que visa a aquisição, interpretação e apresentação dos dados provenientes das três categorias referidas (Dados activos, dados codificados e Dados residentes em áreas não acessíveis), juntamente com sua justaposição contra outras informações disponíveis (por exemplo, transacções de cartões de crédito,

informação bancária, elementos contabilísticos, registos de telefone, Correio Electrónico, documentos diversos, mensagens instantâneas etc.).

A acção da Informática forense não se limita a computadores pessoais e servidores, mas estende-se a todo o tipo de dispositivos que, de algum modo, possam conter informações armazenadas electronicamente (ESI).

Entretanto, persiste ainda uma tendência para considerar que a Informática Forense tem apenas a ver com um tipo específico de criminalidade: o designado “Hi-Tech crime” em que o Computador pode desempenhar quer o papel de “Arma” quer o de “Alvo” do crime.

A verdade é que o, grau de disseminação das novas tecnologias por toda a sociedade actual faz com que muito poucos crimes possam hoje ser cometidos sem recurso à tecnologia, o que faz com que, presentemente, o papel mais comum que o computador desempenha é o de “Repositório” de evidências dos crimes praticados. (É esta a situação mais comum na investigação da criminalidade económica).

Quem realiza este tipo de actividade?

A Informática forense é uma disciplina relativamente recente, mormente no nosso país, onde não existe ainda formação específica nesta área, pelo que, parte dos profissionais que aqui exercem este tipo de atividade fazem-no em grande medida como autodidactas.

O que é possível com a Computação Forense?

Embora a extensão e fiabilidade das informações recolhidas a partir de um exame forense possa variar, adiantam-se alguns exemplos de informações que uma análise pode revelar:

1. Forma e extensão do roubo de dados de propriedade industrial;
2. Timing e extensão de ações de eliminação de arquivos;
3. Se e quando uma pen-drive ou disco rígido externo, foi conectado a um equipamento;

4. Falsificação ou alteração de documentos;
5. Recuperação de e-mail e outro tipo de documentação eletrónica que alguém alegou não existir ou ter sido eliminado;
6. O uso da Internet, as pesquisas on-line e transacções de comércio electrónico (Histórico do uso da Internet);
7. Intrusão e acesso não autorizado a servidores e redes;
8. Manipulação do Relógio e Calendário do sistema;
9. Manipulação de imagens;
10. Retrospectiva, segundo-a-segundo da utilização do sistema.

O que não é possível fazer com a Informática Forense:

Não obstante o grande potencial da Informática Forense, há limites sobre o que pode ser feito com recurso e esta ciência, nomeadamente:

A Computação Forense, geralmente não permite:

1. A Recuperação de qualquer informação que tenha sido completamente sobrescrita por novos dados;
2. Identificar de forma conclusiva se as mãos que manipularam o teclado de um dado sistema foram efetivamente as do utilizador identificado no sistema;
1. Realizar uma análise forense aprofundada, sem acesso ao disco rígido original ou a uma imagem pericial do mesmo;
3. Recuperar dados de uma unidade que sofreu dano físico grave que condicione o respectivo funcionamento;
4. Garantir que a unidade não irá falhar durante o processo de aquisição.

Evidência Digital

A evidência digital não existe por si só, como algo absoluto, trata-se sim, de material que é usado para estabelecer a verdade de um facto particular ou estado de coisas.

A evidência digital não deixa no entanto de ser um tipo de evidência física, embora menos tangível. Ela é composta por campos magnéticos, campos eléctricos e pulsos electrónicos que podem ser recolhidos e analisados recorrendo a técnicas e ferramentas apropriadas.

Características da Evidência Digital

Refira-se, contudo, que a evidência digital possui características muito particulares, que as diferenciam das demais:

- Pode ser duplicada com exactidão, permitindo a preservação da evidência original durante a análise;
- Com métodos apropriados, é relativamente fácil determinar se uma evidência digital foi modificada;
- A evidência digital é extremamente volátil, podendo ser facilmente adulterada durante o processo de análise;
- É difícil de extinguir pois mesmo apagando um arquivo ou formatando um disco rígido, pode ainda ser possível recuperar as informações que este continha.
- É difícil de entender no seu estado puro. Trata-se de campos magnéticos, campos eléctricos e pulsos electrónicos que necessitam de técnicas e ferramentas apropriadas para ser recolhidos e analisados.

Propriedades da Evidência Digital

1. Admissibilidade

   A evidência deve reunir condições para ser aceite como prova em juízo.

   Nos EUA, a lei admite como autênticos os registos gerados pelo computador, desde que os programas que os gerem estejam a funcionar correctamente, considerando, por outro lado, “hearsay evidences” ou seja evidências indirectas equiparadas a “Testemunhos de ouvir dizer” todos os registos em que exista intervenção directa ou indirecta do utilizador. Esta regra admite no entanto excepções, uma das quais relativamente aos documentos relacionados com a actividade económica

das empresas, dado o carácter cíclico e regular com que os documentos são produzidos, acompanhando a dinâmica da actividade. É contudo frequente, que a prova digital assuma no processo um papel de complementaridade, reforçando ou fundamentando outro tipo de prova produzida, (Hoey A. 1996).

No que respeita aos países da União Europeia, de acordo com um estudo realizado em 16 países, publicado em 2006 pelo Journal of Digital Forensic Practice, revelou que em nenhum deles a legislação faz qualquer referência ao termo “Evidência Electrónica”, nem estipulam nas suas normas legais, uma definição específica do que entendem por provas electrónicas, sendo a referência mais directa encontrada no "Police & Criminal Evidence Code" do Reino Unido, referindo:

"*evidências são todas as informações contidas num computador." (Insa, 2006).*

2. Autenticidade

   É indispensável provar que a evidência está relacionada com o incidente de forma relevante, sob pena desta ser considerada inútil.

3. Completude

   A evidência não deve mostrar apenas uma perspectiva do incidente, por exemplo que provem os actos criminalizáveis. É indispensável fornecer igualmente evidências que provem a inocência da vítima.

4. Fiabilidade

   Para que sejam fiáveis, os procedimentos de recolha e análise não podem levantar dúvidas sobre a respectiva autenticidade e veracidade.

5. Legibilidade e Credibilidade

   A forma como se apresentam as evidências ao tribunal, deve permitir a sua fácil compreensão e evidenciar credibilidade.

   Por exemplo, não é aceitável apresentar um dump binário (de uma tabela de base de dados) se os jurados não tem a mais pequena ideia do que isso significa. As evidências devem ser apresentadas num formato standard, compreensível e que possa mostrar a relação com o binário original, de modo a convencer os jurados de que a informação não foi adulterada.

Ciclo da Investigação

Toda a informação relevante deve ser recolhida através de um varrimento meticuloso do dispositivo, para posterior análise, respeitando dois princípios básicos: o da autenticidade, segundo o qual deve ser garantida a origem dos dados, e o da fiabilidade, que assegura que os dados são fiáveis e livres de erros.

Conforme as evidências digitais vão sendo encontradas, devem ser **extraídas, restauradas** quando necessário (caso estejam danificadas ou cifradas), **documentadas e devidamente preservadas**. Em seguida, as evidências encontradas devem ser **correlacionadas,** permitindo a reconstrução dos eventos relacionados com o crime praticado. Muitas vezes na análise das evidências, ao correlacionar e reconstruir os passos seguidos pelo criminoso, promove-se a descoberta de novas informações, formando um ciclo no processo de análise forense.

Áreas de ocultação – Nível Físico

A unidade base de armazenamento de informação num disco rígido é o sector que, na maior parte dos sistemas, corresponde a 512 bytes. Contudo, a generalidade dos sistemas de ficheiros, não utiliza o sector como unidade de atribuição (allocated unit) de espaço em disco, dado o peso excessivo que tal representaria em termos de gestão (gerir uma unidade de disco com 20 GB com base em sectores de 512 bytes, implicaria gerir 40 milhões de sectores específicos).

Para tornar mais eficiente a gestão, os sistemas de ficheiros atribuem os sectores em blocos contíguos, designando estes blocos por clusters.
Um cluster é assim definido como a unidade base de atribuição de espaço de armazenamento ao nível do sistema de ficheiros, sendo constituído por um grupo de sectores consecutivos. O tamanho do cluster (número de sectores que abarca), varia com o dispositivo de armazenamento e é fixado no momento da formatação.

Qualquer unidade de armazenamento de informação necessita de ser previamente formatada para um dado sistema de ficheiros. É assim comum, ter unidades de armazenamento com diversas partições (compartimentos lógicos) formatadas em FAT e/ou NTFS (formatos mais comuns).

Qualquer espaço numa partição que não esteja atribuído a um ficheiro particular não pode ser acedido pelo sistema operativo. Até que esse espaço seja atribuído a um ficheiro, vai permitir ocultar dados.

A criação de uma partição de arranque, obriga a reservar espaço para o Master Boot Record (MBR) no início da unidade. Por força da geometria das unidades de armazenamento esta formatação gera um desperdício de 62 sectores onde poderão ser escondidos dados. A necessidade de criar partições estendidas vai multiplicar estas áreas. (Berghel et all, 2006).

Na criação de partições estendidas que não sejam partições de arranque (boot), vamos ter não 62 mas 63 sectores desperdiçados.

Podemos ter partições removidas de forma premeditada para dissimulação do respectivo conteúdo.

Volume Slack

As partições num disco rígido não utilizam a totalidade do espaço disponível, a área remanescente não pode ser acedida pelo sistema operativo por meios convencionais (por exemplo, através do Windows Explorer). Este espaço desperdiçado é chamado de Volume Slack e pode esconder evidências.

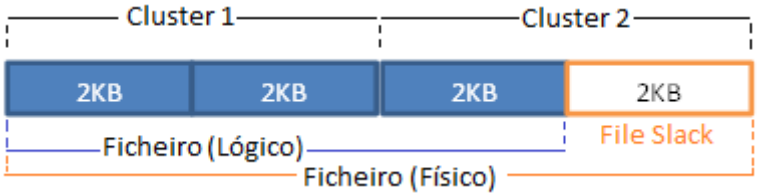
File Slack

Todos os arquivos têm dimensão física e lógica. Em situações normais, a dimensão física é maior do que a dimensão lógica, sendo por vezes igual.

A dimensão física de um ficheiro, é ditada pelo número mínimo de clusters inteiros de que ele necessita. Se por hipótese, a unidade base de atribuição do sistema de ficheiros for de 1 Cluster = 4KB, um ficheiro de 6 KB necessitará de

2 clusters físicos (8 KB), sendo que a sua dimensão lógica apenas irá ocupar 3/4 desse espaço, deixando ¼ (2KB) sem utilização.

A dimensão lógica é o tamanho real do ficheiro que, neste caso, é de 6 KB. A diferença entre as duas dimensões é referenciada como "File Slack" e poderá também ser utilizada para dissimular dados (Berghel et al, 2006).



Fontes de evidências – Nível Lógico

Sistema Operativo faz uso de diversos ficheiros para registo de elementos com relevante interesse para a investigação, como sejam:

“Link Files”, “Swap file”, “Event Logs”, “Recycle Bin”, “Registry”, “Print Spooling”,“Hibernation File”, etc.

Bem como uma estrutura de directórios igualmente relevantes como:

“Recent Folder”, “My Documents”, “Temp Folder”, “Send To Folder”, “Favorites Folder”, “Cookies Folder”, “History Folder”.

Acresce ainda o espaço não atribuído, “Unallocated Clusters”, onde podem igualmente existir importantes evidências.

A análise destes ficheiros e directórios, pode revelar-se determinante, quer para a identificação, quer para confirmação de elementos de prova.