INSTITUTO SUPERIOR TÉCNICO
Departamento de Engenharia Informática
Ciber Segurança Forense
MEIC / METI 2018-2019 – 1º Semestre

Relatório Digital Forense

Autores: Ruben Condesso nº 818969; Miguel Carreiro nº82012; Telma Andrade nº93982

1 Objetivos da investigação

Esta investigação tem como objetivo investigar as ações feitas por *John Mole*. O *John* trabalha há vários anos na empresa *DroneX*, um dos principais fabricantes de tecnologias de *drones*. Devido a uma alteração suspeita de comportamento por parte do *John*, a empresa optou por investigá-lo. *John* tinha acesso privilegiado aos planos de *design* dos novos e revolucionários *drones*, e por conseguinte havia um grande receio que este tivesse roubado informações confidenciais dos *drones* e porventura, vendê-los à concorrência.

De forma a dissipar estes medos, a empresa *DroneX* reuniu uma equipa de auditoria com o objetivo de procurar possíveis evidências de espionagem industrial, depois de ter obtido uma autorização legal. A equipa reuniu um conjunto de arquivos, provenientes de uma *pen drive* que *John* tinha depois de retornar de uma viagem à Alemanha. O objetivo desta investigação passará por analisar esses arquivos e tentar encontrar informação relevante para o problema em causa.

2 Artefactos para análise

Os ficheiros, contidos na *pen drive* do *John*, foram remetidos para um ambiente seguro (sistema *Kali Linux* no nosso PC), para posteriormente serem devidamente analisados.

No total, foram reunidos 8 ficheiros:

- 4 ficheiros de extensão *.png*: ***cathedral.png***, ***oktoberfest.png***, ***street.png*** e ***street.png***;
- 1 ficheiro *.bmp*: ***snow.bmp***;
- 1 ficheiro de extensão *.py*: ***compress.py***;
- 1 ficheiro *.txt*: ***munich.py***;
- 1 ficheiro *.zip*: ***online_banking.zip***.

Os ficheiros *.png* e *.bmp* remetem a que sejam imagens, o *munich.txt* para um ficheiro de texto, o ficheiro *compress.py* aponta para ser um código *python*, e o ficheiro *online_banking.zip* remete para que nele estejam comprimidos outros ficheiros, possivelmente relevantes para a nossa análise. Mas claro, neste tipo de investigações, temos de ter em conta que nada é o que parece ser à primeira vista, e que pode haver informação camuflada em todos os ficheiros e em "todo o lado".

3 Evidências

Neste tópico, iremos relatar quais foram os primeiros passos na nossa investigação tendo em conta o enunciado do projeto, e quais foram as nossas primeiras suspeitas (hipóteses) quando nos deparámos com os ficheiros que tínhamos para análise.

Começamos por copiar o conjunto de arquivos dados para análise para o nosso computador, e de seguida, fazer uma cópia dos mesmos para podemos prosseguir com a investigação, trabalhando sempre na cópia invés do ficheiro original. Para encontrámos as primeiras evidências, fomos abrir os ficheiros para ver qual seria o seu output e começar a tirar as primeiras ilações.

Como primeiro passo, executámos o comando ***md5sum***, para todos os ficheiros, com o intuito de comparar os resultados obtidos desse comando com os valores dados no enunciado, e assim confirmar (ou não) que nenhum

ficheiro tinha sido corrompido ao descarregar para a nossa máquina. Feito isto, verificámos que todos os valores coincidiam, logo os ficheiros tinham sido descarregados corretamente.

O passo seguinte foi abrir os ficheiros de imagem *.png* e o *snow.bmp*. O resultado foram imagens que à primeira vista pareciam normais. Posteriormente, abrimos o ficheiro *munich.txt*, e concluímos que era um ficheiro de texto do *wikipedia*, que por sua vez, também aparentava ser normal. Relativamente ao ficheiro *compress.py*, tentámos abri-lo com um editor de texto mas o resultado foi um conjunto de caracteres e símbolos que não faziam qualquer sentido para nós, logo concluímos que se tratava de um código compilado.

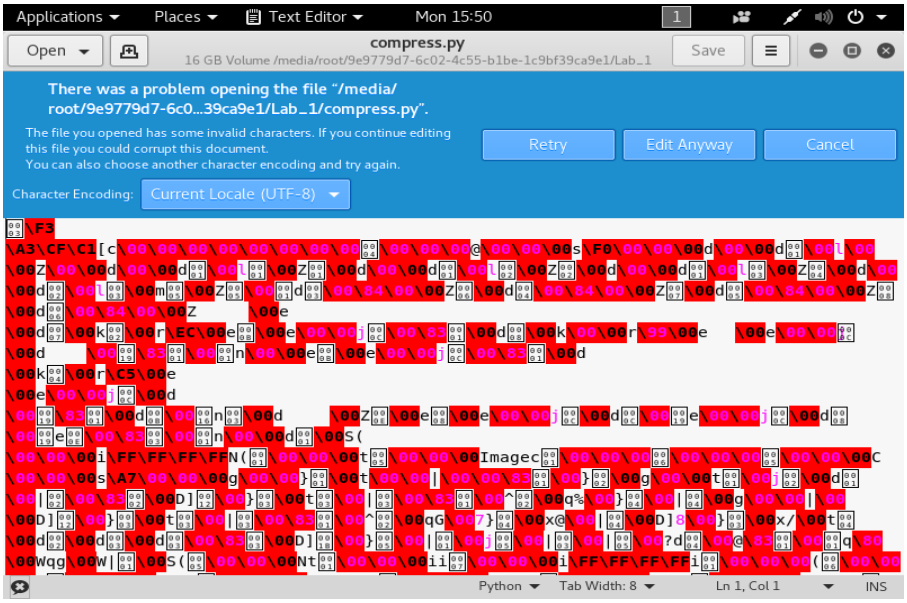

Ilustração 1 - Erro ao abrir o *compress.py*

Finalmente, relativamente ao ficheiro *online_banking.zip*, este encontrava-se protegido por uma *password*, e por conseguinte não conseguimos ver, para já, o seu conteúdo. Nesta altura, todos os ficheiros pareciam à primeira vista normais, ou seja, não pareciam ter sido forjados com alguma informação confidencial. Mas se nos puséssemos na pele de alguém que queria roubar informação confidencial, provavelmente íamos tentar ocultar os nossos passos na melhor forma (e discreta) possível. Posto isto, íamos partir do pressuposto que todos os ficheiros iam ter alguma pista ou alguma informação confidencial, e que todos tinham um propósito, ou seja, que uma pista encontrada num ficheiro poderia levar ao encontro de outra pista que podia estar noutro ficheiro.

Depois de uma primeira análise aos ficheiros e de tentar encontrar uma possível forma de os relacionar, elaborámos as seguintes teorias que irão ser os pontos de partida para a nossa investigação:

1. O *John* escondeu informação confidencial, relativamente aos *drones* da empresa, nas imagens *.png* e/ou na imagem *snow.bmp*. Dado a temática deste trabalho, presumimos que usou um processo de esteganografia, ou seja, escondeu informação confidencial nos *bits* menos significativos das imagem assim as alterações nas imagens não eram visíveis a olho nu. Esta teoria foi também apoiada no facto das imagens terem um tamanho demasiado grande (aparentemente) para a qualidade que apresentavam, o que apontava para o facto de poderem ter informação escondida nas mesmas.

2. O código do ficheiro *compress.py* tenha sido usado para executar o processo de esteganografia referido no parágrafo anterior, e assim havia um propósito específico para este código estar na *pen drive* .

3. O ficheiro *online_banking.zip* deveria ter alguma informação importante, pois doutra forma não estaria protegido por uma *password*, e o ficheiro *munich.txt* parecendo um ficheiro de texto inocente poderia ter um papel importante neste contexto. A nossa terceira teoria foi que a *password* do *online_banking.zip* tivesse entre as palavras do texto, caso contrário, não haveria razão aparente de o ficheiro de texto estar na *pen drive*. Na pior das hipóteses, seria apenas para despistar quem porventura fosse analisar estes ficheiros.

4 Detalhes de examinação

Nesta secção, iremos descrever todos os nossos passos nesta investigação, ou seja, o que fizemos, as ferramentas que usámos, como foi o nosso raciocínio ao longo do tempo e finalmente, o que fomos encontrando.

Começamos por investigar o ficheiro *snow.bmp*. Executámos o comando ***hexdump*** e verificámos que os primeiros dígitos da código hexadecimal correspondia a *4d42*, que remete a ficheiros dessa extensão, portanto,

estávamos mesmo perante um ficheiro *.bmp*. Fizemos, de seguida, o comando *tail* que deu origem à descoberta, no fim do ficheiro de uma mensagem suspeita, numa língua que especulámos ser Alemão, dado que *a priori* sabíamos que o *John* tinha feito uma viagem à Alemanha. Esta mensagem está contida no ficheiro *snow_mensagemOculta*. Usando o *translate* do *google* verificámos que o seguinte significado: " ***Vou enviar-lhe cinco arquivos: (1) planos de drone A, (2) planos de drone B, (3) especificações técnicas, (4) senhas de servidores de arquivos. DroneX*** ", o que nos leva a suspeitar que houve um roubo destes eventuais ficheiros. Com esta descoberta, fizemos a seguinte especulação: ao ler 5 arquivos remete-nos a pensar que poderá ser referente a cada ficheiro *.png* e ao ficheiro *.zip*, indo ao encontro das teorias apresentadas no ponto 3.

Passámos à investigação dos ficheiro *.png*. Executámos o comando ***head*** para cada um e verificamos que os primeiros dígitos do ficheiro dizem *png*, confirmando que estávamos perante ficheiros dessa extensão. De seguida, fizemos o comando ***hexdump*** para cada um, verificámos que os primeiros dígitos em hexadecimal eram ***5089 474e 0a0d 0a1a***, que remetem ao tipo de ficheiro *.png*, estando tudo a bater certo até agora. Tendo em conta a suposição que estes ficheiros de imagem podiam ter sofrido um processo de esteganografia, tendo por base o ficheiro *compress.py*, decidimos então para a investigação desse código, voltando depois a estas imagens.

Executámos o ficheiro *compress.py*, e tal originou a seguinte mensagem no terminal : "*LSB steganography tool: hide files within least significant bits of images" e " the password is optional and must be a number*". Esta frase veio reforçar a nossa teoria de que este código *python* tinha como função esconder informação em forma de texto, nos *bits* menos significativos de uma imagem, ou seja, estávamos perante uma ferramenta de esteganografia. Para investigar este ficheiro, "descompilámos" o seu código usando o programa ***uncompyle2***, e assim conseguimos abrir o código de forma a puder analisá-lo. Feita a sua análise, chegamos à conclusão que o *compress.py* escondia informação, por exemplo um ficheiro de texto, nos 2 *bits* menos significativos de uma imagem. Com o auxílio do programa ***cloacked-pixel-master***, proveniente do *github*, adaptámos algumas das suas funções que nos permitiu criar um programa, *decompress_revers.py*, que tinha como função obter informação que estava escondida nos 2 *bits* menos significativos de uma imagem. Aplicámos o programa às imagens que tínhamos, e conseguimos descobrir os ficheiros que a mensagem escondida na imagem *snow.bmp* referenciava:

- A imagem *oktoberfest.png* deu origem à imagem *oktoberfest_descodificado* onde estava representado um *drone*, onde supusemos que seria o *drone_B*;
- A imagem *wursten.png* deu origem ao ficheiro de texto *wursten_descodificado* que continha as *passwords* dos servidores dos *drones*;
- A imagem *street.png* deu origem à imagem *street_descodificado* que era, a olho nu, igual à imagem original. Tal causo-nos muita estranheza, então voltamos a aplicar o programa *decompress_reverse.py* mas desta vez, na imagem gerada (*street_descodificado*), que por sua vez, deu origem ao ficheiro de texto *street_descodificado2*, onde este continha as especificações dos *drones A* e *B*.

Não havia muitas formas de conseguir esconder alguma informação num ficheiro de texto, ao contrário de um ficheiro de imagem. Posto isto, relativamente ao ficheiro de texto que tínhamos, *munich.txt*, executámos vários comandos, como o *tail*, o *head* e o *hexdump* para tentar encontrar alguma anormalidade, no entanto, todos os *outputs* que tivemos foram dentro do esperado. O único ficheiro que ainda não tínhamos analisado, além deste, era o *online_banking.zip*, que por sua vez estava protegido por uma *password*. Então, como ponto de partida de análise para este ficheiro, seguimos a nossa terceira teoria apresentada anteriormente: a *password* do *.zip* pode estar entre as palavras do ficheiro *munich.txt*. Seguindo este raciocínio, criámos uma cópia do ficheiro *munich.txt*, a que chamámos *dicionario.txt*, com a particularidade que tinha apenas uma palavra por linha.

Assim podemos fazer um ataque por dicionário, usando o programa ***fcrackzip***, para tentar descobrir se alguma das palavras do dicionário correspondia à *password* correta. E o programa obteve uma possível *password*: **Stadelheim**. Desta forma, conseguimos desencriptar o *zip online_banking*. Dentro do *.zip* encontramos dois ficheiros: *online_banking.docx* e *drone-A.bmp*. Ao abrir o primeiro ficheiro, usando o programa *libreoffice*, encontrava-se outra *password*: ***51782***. Tentámos abrir a imagem *drone-A.bmp* mas não conseguimos pois dava o seguinte erro:

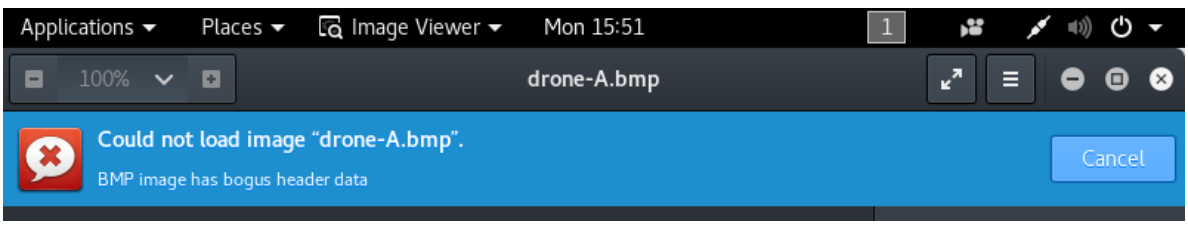

Ilustração 2 - Erro ao abrir a imagem *drone_A.bmp*

Começámos por investigar o código hexadecimal deste ficheiro. No final do mesmo, encontrámos um *timestamp* da imagem: "***tEXtdate:create.2014-12-09T14:18:22+01:00***". Uma pesquisa do excerto "*tEXtdate:create*" no *Google*, levou a uma suspeita de que o formato original da imagem fosse *.png*. Assim, comparámos a assinatura do ficheiro com a assinatura normal da extensão *.png*:

- Assinatura PNG: ***89 50 4E 47 0D 0A 1A 0A***;
- Assinatura ficheiro: **5C 78 30 47 0D 0A 1A 0A**;

Foi possível constatar que os primeiros 6 dígitos hexadecimais da assinatura são diferentes da assinatura normal de um ficheiro *.png*. Assim, procedemos à alteração no ficheiro para que inclua a assinatura correta, bem como à mudança da extensão de *.bmp* para *.png*. Ao reabrir o ficheiro, foi possível visualizar a foto que mostra o esquema de um *drone*, o *drone-A*. Relativamente à *password* encontrada dentro do *.zip*, não conseguimos relacionar com nenhum dos ficheiros encontrados. Com isto, concluímos que tínhamos descoberto todos os segredos relativos a esta investigação, e iremos continuar o relatório com a nossa análise de resultados e a respetiva conclusão.

No que diz respeito a todas as provas encontradas, os respetivos *md5* estão ilustrados na tabela 1:

Ficheiros obtidos	***md5***
compress_reverse.py	9a85bb7268846292o3e9d6c58176288e
drone-A_correto.png	d99f500968d444b5e0a1c9fd1dd69274
oktoberfest_descodificado	cbe4c039f3fa2b312bb95a0964ffba4d
online_banking.docx	b70702822417bd39a7997a0f8c73941f
snow.bmp	a6e56c4d34d9a541b622b74c954c3fc9
street_descodificado	d770b66b4f5833b0be194362f440e494
street_descodificado2	3ba4ca7f05bbf65083360e455fa8ea8a
wursten_descodificado	3cb3f3162e4cf990168d904d3bb300b9

Tabela 1 - *md5's* relativos às provas encontradas

5 Análise de resultados

A nossa suspeita é que o *John Mole* roubou os planos dos *drones* feitos pela empresa *DroneX*, possibilidade que é sustentada pela presença de uma mensagem no ficheiro *snow.bmp*, que explica o conteúdo dos ficheiros retirados indevidamente, explicação essa que vai de encontro ao conteúdo de alguns dos ficheiros encontrados:

- Planos de *drone* A → *drone-A_correto.png*;
- Planos de *drone* B → *oktoberfest_descodificado*;
- Especificações técnicas → *street_descodificado2*;
- Senhas de servidores de arquivos *DroneX* → *wursten_descodificado*.

Uma vez que é dito na mensagem do ficheiro *snow.bmp* que *"Vou enviar-lhe cinco arquivos(...)"*, o *John* poderá ter enviado esses dados a alguém, no entanto, não encontrámos nenhuma prova de que tal tenha sido concretizado efetivamente. O possível envio dos ficheiros, especialmente do ficheiro que contém as senhas dos servidores da empresa e de informações sobre *drones*, e sendo a *DroneX* uma importante fabricante de *drones*, poderá significar que a pessoa que recebeu tais dados terá um interesse muito grande neles, ou seja, é provável que essa pessoa pertença a uma empresa concorrente da *DroneX*.

Em relação a eventuais pagamentos, a única evidência que obtivemos foi o documento *online_banking.docx*, que contém uma *password*, portanto não podemos concluir se houve ou não lugar eventuais recebimentos indevidos por parte do *John*, ou se a password contida nesse ficheiro seja de alguma conta bancária da *DroneX*, sendo que não é de excluir a hipótese de que este ficheiro sirva apenas de distração.

Caso seja dada continuidade à investigação, sugerimos que o computador do *John* seja analisado, incluindo a

sua conta de *e-mail*, de forma a averiguar a existência de mais ficheiros confidenciais da *DroneX* que estejam indevidamente na sua posse ou e-mails comprometedores, bem, como verificar todas as contas bancárias do *John* e da empresa, para averiguar a possibilidade de terem existido transferências monetárias para a conta do mesmo.

6 Conclusões

Esta investigação forense concluiu que *John*, um funcionário da empresa *DroneX*, tinha na sua posse ficheiros sigilosos da empresa em causa, onde estes sofreram técnicas de ocultação para que não fossem detetados à primeira vista. A equipa forense indicou algumas pistas para estudo, caso exista intenção de prosseguir com a investigação e encontrar novas provas, e/ou confirmar algumas hipóteses levantadas neste relatório. A empresa poderá usar este relatório para acionar mecanismos judiciais contra o funcionário, e rever os seus procedimentos internos de segurança, de forma a evitar novas fugas de informação.