TOMASZ ŁAPAJ

Odzyskiwanie danych z pliku wymiany

Stopień trudności

W artykule zostały przedstawione dwie podstawowe techniki analizy zawartości pliku wymiany, skopiowanego z sytemu Windows XP. Przedstawiono, jak za pomocą skalpela wydobyć zapisane tam pliki oraz możliwości, jakie daje nam wyszukiwanie słów kluczowych przy użyciu edytora szesnastkowego Bless.

Plik wymiany jest swego rodzaju przedłużeniem pamięci operacyjnej. Dane, które są w danej chwili niepotrzebne, system operacyjny usuwa z pamięci RAM i zapisuje w pliku wymiany. Zjawisko to jest szczególnie intensywne, gdy RAM-u jest za mało i objawia się, wszystkim nam dobrze znanym, mieleniem dysku. Z punktu widzenia informatyki śledczej wniosek nasuwa się sam – jeżeli jakiekolwiek dane były zawarte w pamięci operacyjnej (obrazy, dokumenty, poczta elektroniczna, hasła) to istnieje prawdopodobieństwo, że mogły one zostać skopiowane do pliku wymiany. Co ciekawe, jeżeli nasz podejrzany jest zapobiegliwy i sprytny, będzie zacierał za sobą ślady – wyczyści prywatne dane z Firefoksa, opróżni kosz, użyje Erasera (*http://eraser.heidi.ie/*) do wymazania usuniętych plików i CleanAfterMe (*www.nirsoft.net*) do wymazania przeróżnych śladów aktywności w systemie. Żeby wyczyścić plik wymiany, trzeba dużo więcej pracy. Jest to plik chroniony przez system operacyjny, dlatego trzeba uruchomić Linuksa z płyty CD i wymazać plik ręcznie. Trzeba być naprawdę zaawansowanym użytkownikiem, żeby w ogóle wpaść na taki pomysł.

Scenariusz

Analizie poddano *pagefile.sys* skopiowany z systemu Windows XP SP3, zainstalowany na Sun VirtualBox.

System został wyposażony w 256 MB pamięci RAM. Nie jest to dużo, ale i tak system raportował 60 MB wolnej pamięci, więc stronicowanie nie było wymuszone.

Zainstalowano następujące aplikacje: Firefox 3.5, OpenOffice.org 3, a następnie:

- zarejestrowano się i wysłano kilka e-maili, korzystając z poczty Gmail,
- wyszukano na stronie www.google.com i przejrzano zdjęcia planet,
- stworzono i zapisano na pulpicie dokument MS Word.

Do testów użyto Debiana 5 Lenny, weryfikując część rezultatów w Ubuntu 9.04.

Scalpel – instalacja

Strona domowa aktualnej wersji 1.6: *http://www.digitalforensicssolutions.com/Scalpel/*.

Instalacja nie powinna przysporzyć problemów. W Debianie Lenny po prostu wpisano:

```
#  aptitude install scalpel.
```

Dostępny jest również pakiet rpm i zapewne w innych dystrybucjach instalacja nie przysporzy nikomu wielu problemów. Jeżeli instalujemy ze źródeł, musimy je najpierw rozpakować:

```
$ tar -xvzf scalpel-1.60.tar.gz.
```

Z ARTYKUŁU DOWIESZ SIĘ

jak odzyskiwać pliki z pamięci wymiany,

jak analizować jej zawartość,

jak chronić swoją prywatność przed powyższymi technikami.

CO POWINIENEŚ WIEDZIEĆ

jak używać Linuksa – instalacja programów, podstawy linii poleceń,

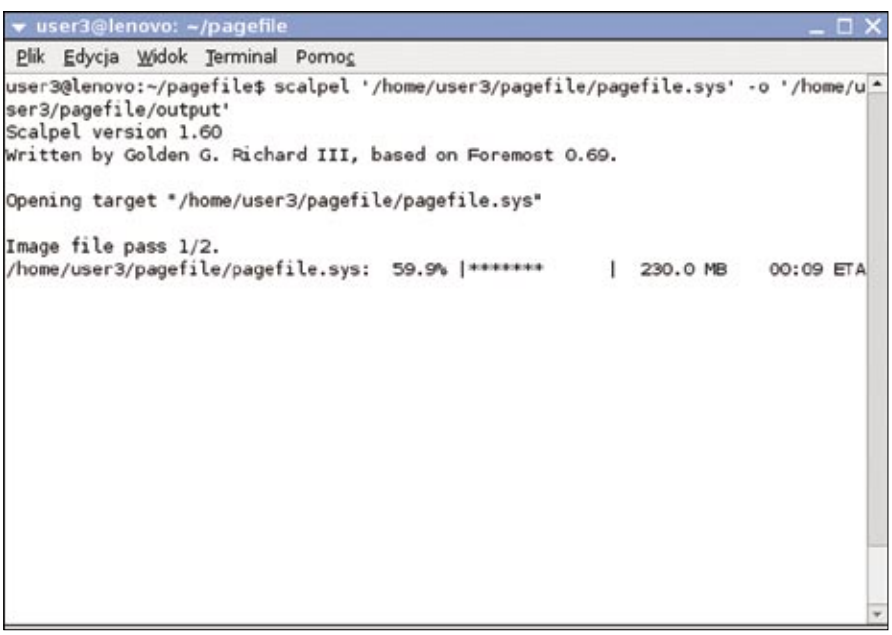
jak używać Windows.

Po przejściu do rozpakowanego katalogu źródła kompilujemy poleceniem: `$ make`, a następnie orientujemy się, że polecenie `make install` nie zostało zaimplementowane, więc jeżeli przeszkadza nam uruchamianie programu z właśnie skompilowanej lokalizacji przenosimy plik Scalpel do */usr/local/bin*, a plik podręcznika *scalpel.1* przenosimy do */usr/local/man/man1*. Zapraszamy do zapoznania się z treścią README. Nie ma gwarancji, że wyżej wymienione katalogi będą prawidłowe na wszystkich systemach, więc instalacja ze źródeł może wymagać dodatkowej wiedzy.

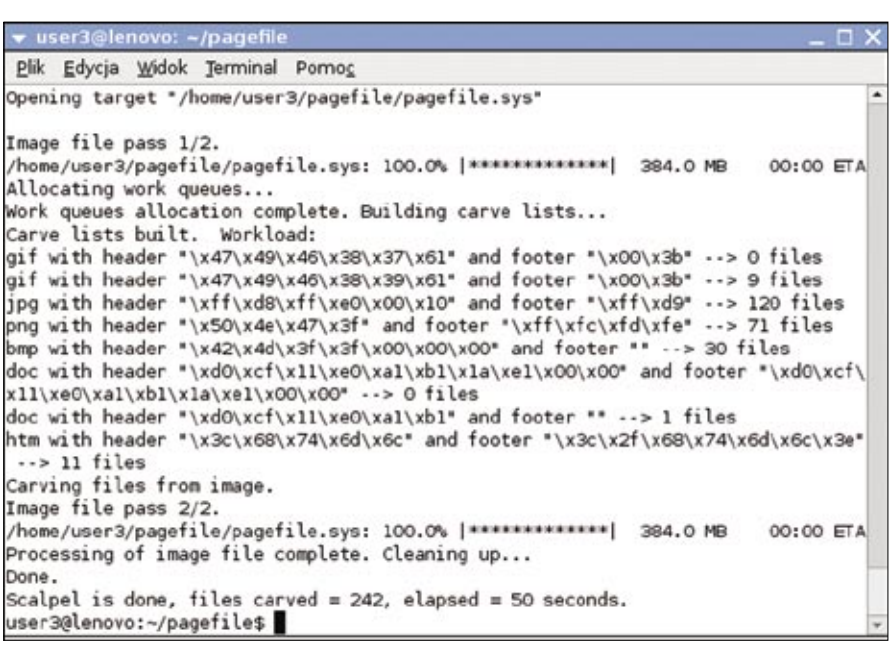
Sygnatury plików

Większość plików posiada sygnatury. Sygnatura określa typ pliku. Na przykład pliki graficzne *.jpg* mają sygnaturę *ff d8 ff e0 00 10*. Łatwo to sprawdzić – wystarczy otworzyć dowolny plik *.jpg* w edytorze szesnastkowym i zobaczyć od jakich znaków się zaczyna. To właśnie dzięki sygnaturom Linux rozpoznaje pliki i wskazuje aplikację, która powinna je otworzyć. W przeciwieństwie do Windows rozszerzenia nie są tutaj potrzebne.

Scalpel używa sygnatur w celu zlokalizowania plików na wskazanych partycjach, dyskach, obrazach dysków lub plikach. Po znalezieniu sygnatury początku pliku (nagłówka) Scalpel szuka sygnatury końca (stopki) i plik jest gotowy do wycięcia. Niestety nie zawsze jest tak łatwo – pliki często są częściowo nadpisane, nie wszystkie typy plików posiadają sygnaturę końca, ponadto często sygnatury są zbyt krótkie i generują dużą ilość fałszywych trafień. Scalpel korzysta ze złożonych algorytmów w celu ominięcia tych problemów i rezultaty są całkiem dobre.



Rysunek 1. *Scalpel rozpoczyna operację odzyskiwania plików*



Rysunek 2. *Udało się odnaleźć 252 pliki*

Plik konfiguracyjny

Ścieżka do pliku konfiguracyjnego jest następująca: */etc/scalpel/scalpel.conf*. Każda linia opisuje parametry wyszukiwania jednego typu pliku, a raczej jednej sygnatury. Spójrzmy na linię, opisującą archiwum zip:

```
# zip     y     10000000     PK\x03\x04
                             \x3c\xac
```

Hash oznacza komentarz, więc trzeba go usunąć, jeżeli chcemy żeby Scalpel szukał plików tego typu. *y* oznacza, że sygnatura rozróżnia duże i małe litery. 10000000 – to maksymalna wielkość pliku w bajtach. *PK\x03\x04* – sygnatura początku pliku. PK to tekst ASCII, za którym mamy dwa bajty w zapisie szesnastkowym – *03* i *04*. *\x3c\xac* to sygnatura końca pliku.

Scalpel w akcji

Celem analizy jest odzyskanie plików graficznych, dokumentu MS Word oraz przeglądanych stron internetowych. W pliku konfiguracyjnym należy więc usunąć

Rysunek 3. *Odzyskany plik graficzny*

Rysunek 4. *Odzyskany plik graficzny*

komentarze do następujących typów plików:

- *bmp*
- *doc*
- *gif*
- *htm*
- *jpg*
- *png*

Scalpel do działania nie wymaga wielu parametrów. Wystarczy wpisać: `scalpel plik_do_analizy`. Używając opcji `-o` możemy wskazać katalog docelowy dla odzyskanych plików:

```
$ scalpel '/home/user3/pagefile/
   pagefile.sys' -o '/home/user3/
                     pagefile/output'
```

Jeżeli program zainstalowany został ze źródeł lub z innego powodu plik konfiguracyjny nie znajduje się w domyślnej lokalizacji, skorzystajmy z opcji `-c ścieżka_do_pliku_konfiguracyjnego`. Przykładowy plik konfiguracyjny spakowany jest wraz ze źródłami.

Scalpel odczyta plik konfiguracyjny, przeanalizuje plik źródłowy i zapisze odzyskane pliki w katalogu docelowym.

Wyniki – Scalpel

Wycięte pliki Scalpel kopiuje do katalogu docelowego. Scalpel odzyskał 252 pliki. Bez trudu otworzono 93 pliki graficzne oraz 2 dokumenty HTML. Pozostałe to fałszywe trafienia lub pliki wymagające wnikliwszej analizy. Zmieniając ustawienia Scalpela pamiętać zawsze należy o zachowaniu równowagi pomiędzy fałszywymi trafieniami, a ryzykiem pominięcia ważnego materiału dowodowego. W naszym przypadku fałszywe trafienia to nie problem, gdyż można je szybko wyeliminować, przeglądając zdjęcia w menedżerze plików z widokiem galerii, takim jak Nautilus.

Rysunek 5. *Plik .jpg odzyskany częściowo*

Dwa odzyskane pliki HTML związane są z zakładaniem konta Gmail. 93 pliki jpg to w większości pliki przedstawiające planety. Udało nam się udowodnić że rzeczywiście przeglądałem te zdjęcia!

Wyniki – Bless

Nie udało się niestety odzyskać dokumentu MS Word ani korespondencji Gmail, zachowanej w postaci pliku HTML.

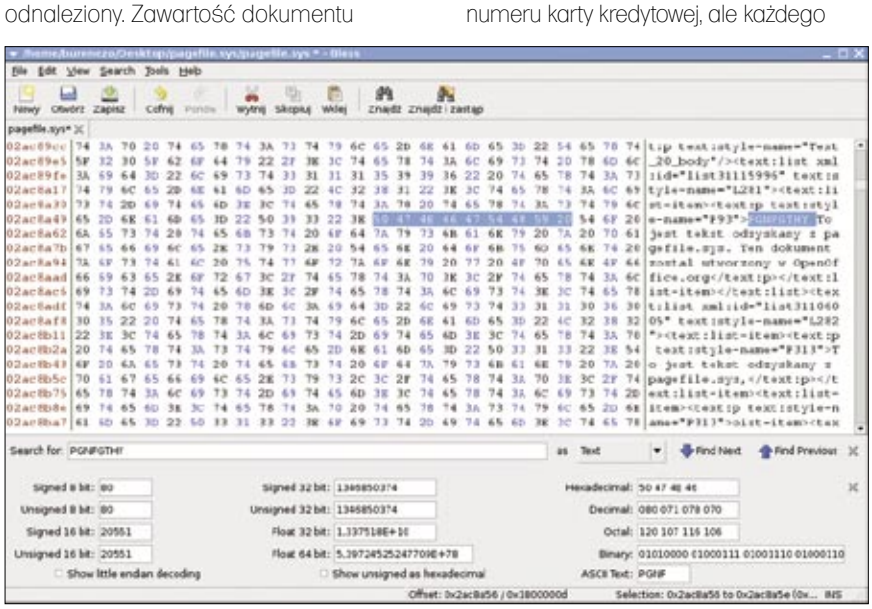
Nie traćmy nadziei. Z pomocą przyjdzie nam edytor szesnastkowy Bless, który na Debianie instalujemy za pomocą polecenia: `# aptitude install Bless` (lub instalacja zgodna z używaną dystrybucją, bądź upodobaniami). Przed otwarciem *pagefile.sys* zmodyfikujmy tak prawa dostępu, żeby uniemożliwić sobie przypadkowego nadpisania pliku. Możemy zacząć poszukiwania. Ułatwiłem sobie zadanie i w pliku MS Word umieściłem słowo kluczowe PGNFGTHY. Techniki poszukiwania bliżej nieskonkretyzowanych danych opisano pokrótce w dalszej części artykułu. Należy pamiętać, aby szukać ciągu tekstowego, a nie szesnastkowego – w przeciwnym wypadku Bless zwróci błąd lub będzie szukał nieprawidłowych danych. Po kilku sekundach tekst został odnaleziony. Zawartość dokumentu MS Word została skopiowana do pliku wymiany!

Kolejny krok to konwersacja e-mail. Również w tym wypadku w poszukiwaniach pomaga wiedza, co do treści korespondencji. Po kilku sekundach odnaleziony został jeden z e-maili. Dowodem jest Rysunek 7.

Czas na bonus. Do pola wyszukiwania w Bless wpisano hasło, którego użyto do zalogowania się w Gmail. Po kliknięciu wyszukiwania hasło zostało znalezione. I to w kilku miejscach!

Igła w stogu siana

Sprawa z Bless wygląda nieco trudniej, jeżeli nie wiemy czego szukać. Jeżeli sprawa jest wagi państwowej i mamy nieograniczone środki, polecam przejrzenie pamięci wymiany, strona po stronie w poszukiwaniu interesujących informacji. Jeśli przyjmiemy, że strona zawiera 2048 znaków, to w 1 GB pliku wymiany będzie zaledwie 500 tys. stron. Zatrudniamy 1000 analityków i sprawę mamy załatwioną w 1 dzień. Jako, że taka opcja będzie raczej niemożliwa do zrealizowania, musimy przede wszystkim przygotować przemyślany zestaw słów kluczowych. Jeżeli szukamy poczty elektronicznej, dobrym startem będą słowa witam, pozdrawiam, imiona i nazwiska związane ze sprawą, słowa, które mogły być zawarte w korespondencji, adresy email. Bardzo pomocne będą wyrażenia regularne (polecenie `grep`), dzięki którym możemy poszukiwać nie tylko konkretnego numeru karty kredytowej, ale każdego



Rysunek 6. *Bless. Fragment dokumentu MS Word*

ciągu 16 cyfr, w grupach po 4 oddzielonych spacją lub myślnikami.

Kolejna technika, jaką możemy zastosować, to polecenie `strings`.

```
$strings pagefile.sys > slowa_w_
                        pliku.txt
```

`Strings` analizuje pliki binarne w poszukiwaniu znaków drukowalnych. Jeżeli znajdzie co najmniej 4 znaki drukowalne zakończone znakiem niedrukowanym, uzna dany ciąg za słowo i wyświetli w konsoli lub jak w powyższym przykładzie zapisze do pliku *slowa_w_pliku.txt*. Danych wyjściowych może być ciągle dużo, będzie tez sporo fałszywych trafień, ale ich analiza w przeciwieństwie do przeglądania pliku wymiany strona za stroną staje się możliwa. Jako ciekawostkę podam, że dane wyjściowe polecenia strings są dobrym pomysłem na atak słownikowy. Zwłaszcza, jeżeli jako danych wejściowych użyjemy całego dysku podejrzanego.

Jak się bronić

Jak widać plik wymiany może zawierać sporo informacji o aktywności użytkownika. Co ciekawe, nawet użytkownicy Linuksów uruchamianych bezpośrednio z płyty, nie mogą czuć się bezpiecznie. Jeżeli system znajdzie partycje wymiany (swap) w większości przypadków ja zamontuje i może tam zapisać prywatne informacje.

Najlepszym sposobem na obronę jest całkowite wyłączenie pliku wymiany. Możemy sobie na to pozwolić oczywiście tylko w wypadku, gdy nasz system wyposażony jest w wystarczającą ilość pamięci operacyjnej. Jak dużo to wystarczająco? Zależy to od wielu czynników, takich jak rodzaj systemu operacyjnego, zainstalowanych aplikacji, wielkości otwieranych dokumentów. Dla przykładu typowy Linux lub Windows XP z 1 GB RAM powinien sobie doskonale poradzić, co do Visty są już jednak spore wątpliwości. Najlepszym sposobem na sprawdzenie konkretnego systemu, jest uruchomienie kilku pamięciożernych aplikacji i sprawdzenie ilości wolnej pamięci. Jeżeli jest jej dalej dużo, możemy wyłączyć plik wymiany. Drugi sposób to całkowite wymazanie danych zawartych w pliku wymiany, czyli nadpisanie ich ciągiem zer lub losowych wartości.

Windows

W Windows XP ustawienia pliku wymiany są porządnie zakopane i trzeba się namęczyć, żeby się do nich doklikać. Zaczynamy od prawego kliknięcia na Mój Komputer, następnie wybieramy właściwości (lub alternatywnie w Panelu Sterowania klikamy na System), dalej Zaawansowane, Wydajność systemu, Ustawienia, Zaawansowane,Pamięć wirtualna, Zmień. Uff, udało się. Teraz wystarczy wybrać dysk, na którym zapisany jest plik wymiany, wybrać Brak pliku wymiany z listy opcji i kliknąć na Ustaw.

Możemy również ustawić system operacyjny, aby wymazywał zawartość pliku wymiany w trakcie zamykania systemu. W rejestrze Windows XP oraz Vista należy upewnić się że wartość `ClearPageFileAtShutdown` wynosi 1:

```
HKEY_LOCAL_MACHINE\SYSTEM\
   CurrentControlSet\Control\Session
      Manager\Memory Management \
         ClearPageFileAtShutdown
```

System nadpisze wszystkie dane, nieużywane przez Windows w trakcie jego zamykania. Oznacza to, że część pliku wymiany nie zostanie nadpisana, ale będzie to dotyczyć raczej danych systemowych, niż użytkownika. Do edycji rejestru używamy polecenia `regedit`. Oczywiście, jak zawsze w przypadku dokonywania zmian w rejestrze nie zalecam ostrożności. System może się już nie uruchomić i istnieje ryzyko utraty danych, ale jego naprawa to przecież najlepsza metoda nauki.

Linux

Jądra Linuksa odczytuje konfigurację pamięci wymiany z pliku */etc/fstab*. Za pomocą polecenia `swapon` oraz `swapoff` możemy włączać i wyłączać pamięć wymiany na uruchomionym systemie oraz sprawdzać jej status. Polecenie `mkswap` formatuje partycje lub dowolny plik jako plik wymiany. Polecenia `dd` możemy użyć do nadpisania danych. Na początek sprawdźmy status pamięci wymiany:

```
#swapon -s
Filename  Type  Size  Used  Priority
/dev/hda5  partition  409616  84  -1
```

Teraz możemy ją zdeaktywować:

```
#swapoff /dev/hda5
```

Ponownie sprawdzamy status:

```
#swapon -s
Filename  Type  Size  Used  Priority
```

Jak widać partycja wymiany została wyłączona. (Kilka dodatkowych poleceń i parametrów: `swapoon /dev/sda5`, aby

```
07b24776 69 6D 65 73 2C 20 73 65 72 69 66 22 3E 0D 0A 3C 44 imes, serif">..<D
07b24787 49 56 3E 57 69 74 61 6D 2C 20 75 20 6D 6E 69 65 20 IV>Witam, u mnie
07b24798 77 73 7A 79 73 74 6B 6F 20 77 20 70 6F 72 7A 61 64 wszystko w porzad
07b247a9 6B 75 2E 20 50 72 7A 65 67 6C 61 64 61 6D 20 7A 64 ku. Przegladam zd
07b247ba 6A 65 63 69 61 20 70 6C 61 6E 65 74 2E 20 53 61 20 jecia planet. Sa
07b247cb 6E 61 70 72 61 77 64 65 20 6E 69 65 7A 6C 65 2E 20 naprawde niezle.
07b247dc 4D 61 6D 20 6E 61 64 7A 69 65 6A 65 20 7A 65 20 6D Mam nadzieje ze m
07b247ed 6E 69 65 20 6E 69 6B 74 20 6E 69 65 20 70 72 7A 79 nie nikt nie przy
07b247fe 6C 61 70 69 65 20 62 6F 20 62 65 64 61 20 70 72 6F lapie bo beda pro
07b2480f 62 6C 65 6D 79 2E 2E 2E 20 52 6F 6D 65 6B 3C 2F 44 blemy... Romek</D
07b24820 49 56 3E 0D 0A 3C 44 49 56 3E 26 6E 62 73 70 3B 3C IV>..<DIV> <
07b24831 2F 44 49 56 3E 0D 0A 3C 44 49 56 3E 4F 62 63 79 26 /DIV>..<DIV>Obcy&
07b24842 6E 62 73 70 3B 3C 42 52 3E 3C 2F 44 49 56 3E 0D 0A nbsp;<BR></DIV>..
```

Rysunek 7. *Fragment e-mail*

```
33 63 38 34 39 62 ID=4986cb963c849b
33 33 30 33 32 3A a7:TM=1247333032:
33 32 3A 53 3D 79 LM=1247333032:S=y
38 41 73 66 3B 20 H_w7gQsNX5hXAsf;
39 4C 5F 52 54 54 TZ=420; GMAIL_RTT
31 3A 20 6E 6F 2D =102..Pragma: no-
65 2D 43 6F 6E 74 cache..Cache-Cont
38 65 0D 0A 0D 0A rol: no-cache....
38 61 73 6C 6F 36 Passwd=mojeHaslo6
73 6D 69 74 68 31 &Email=johnsmith1
```

Rysunek 8. *Zwykle używam bardziej wyszukanych haseł niż "mojeHaslo6"*

```
6F 6D 26 k=googlemail.com&
6C 6F 36 Passwd=mojeHaslo6
6F 6A 65 &PasswdAgain=moje
65 6E 74 Haslo6&Persistent
68 6F 77 Cookie=yes&rmShow
68 6B 3D n=1&smhck=1&nshk=
61 74 2B 1&selection=What+
2B 70 68 was+your+first+ph
```

Rysunek 9. *Hasło wpisane w trakcie rejestracji w Gmail*

aktywować ją ponownie, `swapoff -a`, aby zdeaktywować wszystkie partycje i pliki wymiany wyliczone w pliku */etc/fstab*, *swapon -a*, aby je aktywować). Zostanie ona aktywowana ponownie przy następnym uruchomieniu systemu. Aby tego uniknąć musimy zmodyfikować plik */etc/fstab*. Użyjemy do tego edytora tekstowego gedit, jednak do wpisania jednego znaku każdy inny jest równie dobry:

```
# gedit /etc/fstab
```

Wystarczy wstawić znak komentarza `#` na początku linii odpowiedzialnej za zamontowanie swapu, czyli tej w której znajdziemy słowo swap:

```
#/dev/hda5  none  swap  sw  0  0
```

Po ponownym uruchomieniu systemu ponownie weryfikujemy czy partycja wymiany została zamontowana:

```
$swapon -s
```

Nadpisanie pliku wymiany w Linuksie również ogranicza się do kilku prostych poleceń – najpierw deaktywujemy pamięć wymiany, tak jak powyżej:

```
# swapoff /dev/hda5
```

Następnie, korzystając z polecenia dd nadpisujemy zdeaktywowaną partycje ciągiem zer:

```
# dd if=/dev/zero of=/dev/hda5
```

W zależności od rozmiaru pliku wymiany oraz szybkości i aktualnego obciążenia dysku może to zająć trochę czasu. Przyjmijmy ze nasz dysk zapisuje dane z prędkością 50 MB/s. Nadpisanie 1 GB partycji wymiany zajmie więc 20 sekund.

Polecenie `dd` binarnie (bit za bitem) kopiuje zawartość plików: parametr `if` to *input file*, czyli plik wejściowy, `of` to *output file*, czyli plik wyjściowy, `/dev/zero` to urządzenie wysyłające ciąg zer, `/dev/hda5` to nasza partycja wymiany. Podsumowując dd wypełni plik wymiany danymi odczytanymi z urządzenia `/dev/zero`. Jeżeli do nadpisania chcemy użyć danych losowych, skorzystajmy z urządzenia `/dev/random` zamiast `/dev/zero`. Nie, nie ma urządzenia `/dev/one`.

Jeżeli ponownie chcemy użyć partycji wymiany należy ja sformatować:

```
# mkswap /dev/hda5
```

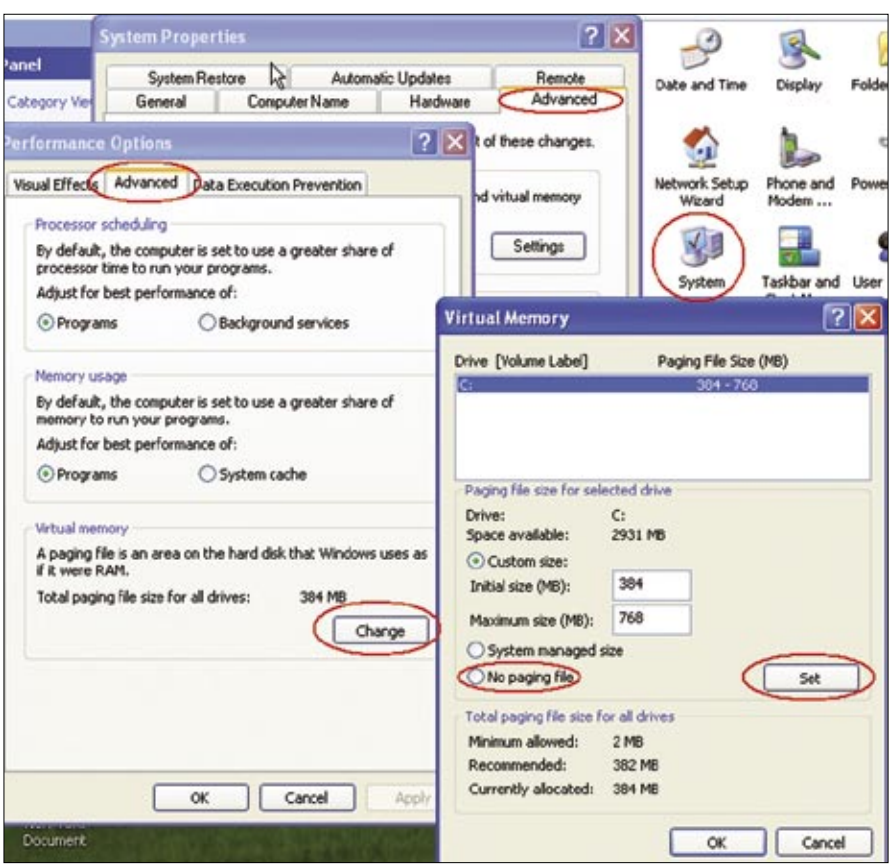
I aktywować:

```
# swapon /dev/hda5
```

Skrypt zawierający powyższe polecenia (swapof, dd, mkswap) możemy oczywiście dodać do listy skryptów uruchamianych wraz z zamknięciem systemu.

Podsumowanie

Jak widać plik wymiany stanowi źródło informacji, które nie może być pominięte w trakcie analizy systemu. Zawartość pliku wymiany jest dynamiczna. System operacyjny stale zapisuje tam nowe dane, osuwając przy tym poprzednie. Nigdy w pełni nie będziemy mogli przewidzieć, co tam znajdziemy, jednak jak widać odzyskane dane mogą być niezwykle interesujące. Analizie poddany został wprawdzie Windows XP, jednak w trakcie pisania artykułu przyjrzałem się również bliżej pamięci wymiany Windows Vista, Windows 7 i Linuksa – we wszystkich przypadkach można było odzyskać pliki, najczęściej graficzne oraz dane tekstowe.



Rysunek 10. *Kolejne kroki do okna odpowiedzialnego za zarządzanie pamięcią wymiany w Windows XP*

Jak zdobyć *pagefile.sys*? Plik wymiany w Windows XP nazwany został *pagefile.sys* i domyślnie znajdziemy go bezpośrednio na dysku C. Jest on chroniony przez Windows XP i nie da się go po prostu skopiować z uruchomionego systemu. Sposobem na obejście tego problemu może być zamontowanie partycji w Linuksie, skopiowanie pliku wymiany na zewnętrzny nośnik. Darmowy, chociaż nie wolny, program FTK Imager (*www.accessdata.com*) jest w stanie skopiować pliki systemowe, nawet z uruchomionego systemu Windows, co czasem może znacznie ułatwić pracę.

Tomasz Łapaj
Autor pracuje jako digital forensics analyst. Posiada certyfikaty ISFCE CCE, ENCE, ACE.
Kontakt z autorem: *t@fr.pl*