BROWSWER SECRETS UNVEILED

By Chetan Gupta

Everyday millions of people surf the web using popular web browsers such as Microsoft Internet Explorer (IE) or any one from the Firefox/Mozilla/Netscape family. A very important step in computer forensics is investigating the web usage of the suspect. This information is useful in everything from examining company policy violation to detecting corporate espionage. Examining a suspect's web browsing history could provide critical clues to solving the case.

Each of these browsers saves the web browsing activity in their own unique formats. The Internet activity data related to a specific browser could be found in different locations according to the Operating System used by the suspect. In this article, we look at the various tools and techniques available for investigating one of the most widely used browsers: Internet Explorer.

Investigating Microsoft's Internet Explorer (IE)

IE has been the default browser for Windows OS since Windows 95. IE has evolved to closely interact with the OS which opens various avenues for the forensics investigation. Even if the suspect deletes the history or clears the cache, the information can still be recovered to a great extent.

IE has three separate logging facilities that can be used to reconstruct the suspect's web browsing activities. The logged information includes

1. History of Visited URLs

The browser history creates an audit trail that shows the user's browsing patterns on the web. In Windows XP, the files are stored in the following locations:

- C:\Documents and Settings\profile_name\Local Settings\History\History.IE5\index.dat
- C:\Documents and Settings\profile_name\Local Settings\History\History.IE5\MSHist01XXXXXXXXXXXXXXXX\index.dat.

  The MSHist files are of two types:

  - Daily MSHist files (ie, MSHist011998100819981009index.dat: one day between yyyymmddyyyymmdd dates)
  - Weekly MSHist files (ie, MSHist011998092819981005\index.dat: one week between yyyymmddyyyymmdd dates)

2. Cookies

Cookies are small text files that the suspect's browser accepted while browsing the web. The information held in these files can be invaluable to forensics. In Windows XP, the files are stored in the following location:

C:\Documents and Settings\profile_name\Cookies\index.dat

This audit trail in the index.dat is useful in locating cookies that are intentionally misnamed or obfuscated.

3. Temporary Internet Files

The Temporary Internet Files folder contains a copy of the files that constituted the web pages the suspect visited. The next time the web page is requested, only the sections of the page that have changed would be downloaded. In the cache directory, an index.dat maps the pages on the websites to files and directories in the cache. In Windows XP, the files are stored in the following locations:

- C:\Documents and Settings\profile_name\Local Settings\Temporary Internet Files\Content.IE5\index.dat
- C:\Documents and Settings\profile_name\UserData\index.dat

The index.dat in the UserData folder holds information about automatic Windows accesses to the Internet such as Windows updates. Each of these facilities contains an important file called index.dat which is primarily a catalog file for the saved information. The index.dat files can catalog the URLs visited, the cookies accepted by the browser and locally cached files linked to the visited URLs.

Reconstructing IE History

The first place an investigator would start his investigation is by reviewing the URL history of the browser used by the suspect. The history files are located in the C:\Documents and Settings\profile_name\Local Settings\History\History.IE5 directory. The directory History.IE5 when viewed in the browser shows the following files:

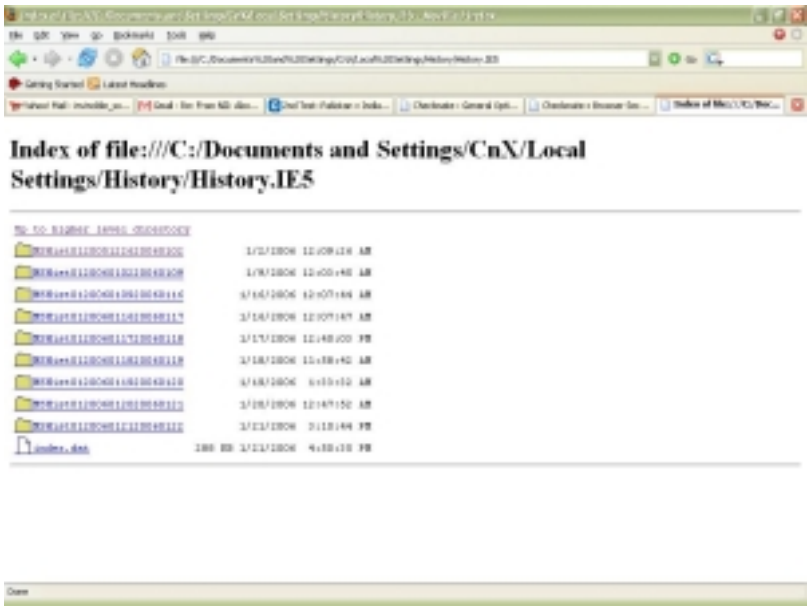

Figure 1: Browser View of IE5 History Files

The index.dat file shown above contains the history information for the date of last access. There are different directories starting with MSHist01 followed by a string of numbers. This string of numbers actually represents the dates for which the file holds the history information. For example, the number 2005020320050210 means that the file holds the history information

for the period between 2005-02-03 and 2005-02-10. Each MSHist01XXXXXXXXXXXXXXXX directory contains an index.dat file which contains history records for a specific period.

There are two ways of accessing the information contained in the history files:

1. The easiest way to see the contents of browser history is by opening the IE browser and pressing Ctrl+H or select View tab -> Explorer Bar -> History

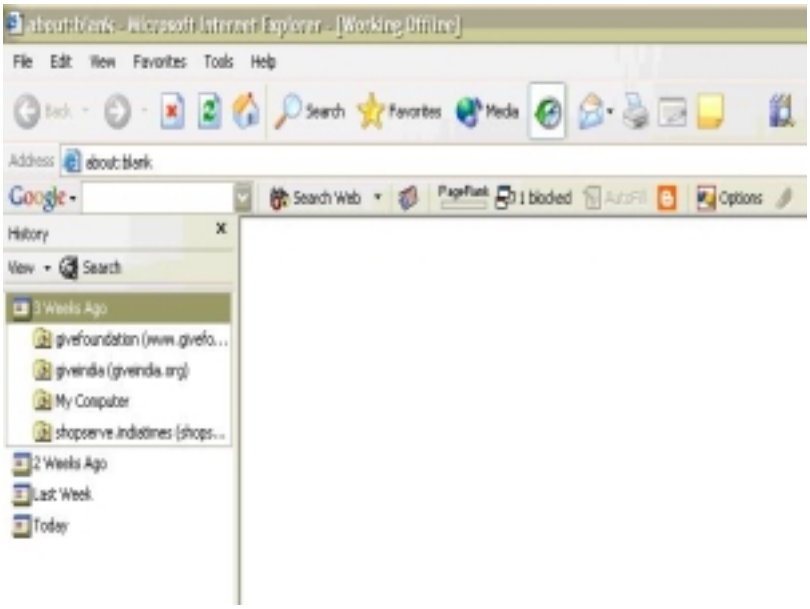

Figure 2: IE History Files

2. Using a tool such as Pasco or any commercial tool such as Encase

Tool: Pasco

Pasco is an open source Web Browsing Investigation tool. Pasco examines the index.dat files which are populated when a suspect browses the Internet. Here is a sample output in Excel format generated by running Pasco on the index.dat file:

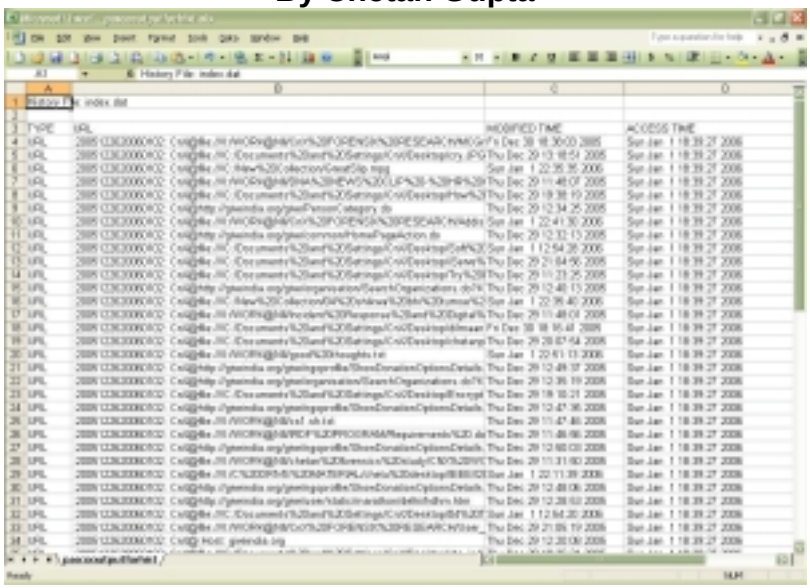

Figure 3: 'Index.dat' Entries

Reconstructing Browser Cache

To speed up browsing of the Internet, IE caches most of the pages you visit on your hard drive. This is good news for Forensics investigators! In order to avoid confusion and collisions among the similarly named files (eg. index.html), Microsoft uses a naming system which allows the files to be stored in different randomly named directories and the catalog in the index.dat file allows the mapping between the filename and the associated directory where it can be found. Here is a sample output in excel format generated by running Pasco on the index.dat file

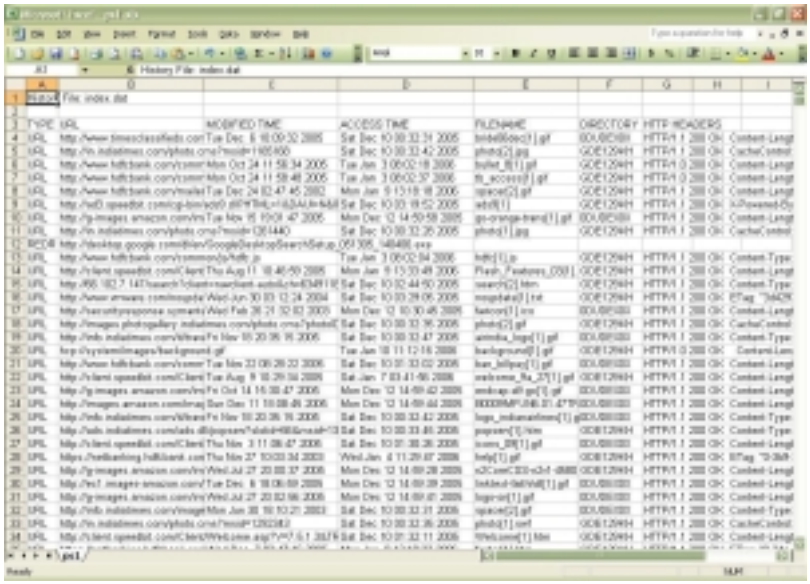

Figure 4: 'Index.dat' Entries

Let us look at a sample output from the index.dat file obtained by running the following command at the command prompt: C:\Documents and Settings\CnX\Local Settings\Temporary Internet Files\Content.IE5 -> pasco.exe index.dat -> c:\pasco_output.doc
We will have a look at the first entry and try to interpret it.

History File: index.dat

File Structure

TYPE URL MODIFIED TIME ACCESS TIME FILENAME DIRECTORY HTTP HEADERS

URL http://www.timesclassifieds.com/tc/timesmatri/images/bride06dec.gif Tue Dec 6 10:09:32 2005 Sat Dec 10 00:32:31 2005 bride06dec[1].gif 0DUDEX0X HTTP/1.1 200 OK Content-Length: 1758 Keep-Alive: timeout=10, max=100 Content-Type: image/gif Content-Language: en-US ~U:cnx

Interpretation

1. The type is 'URL'
2. The URL visited was http://www.timesclassifieds.com
3. The date on which this entry was added to the cache is 6th December 2005 at 10:09:32 am in GMT format (Local time (India) is +5:30 GMT which is 03:39:32PM)
4. The date on which this entry was last accessed from the cache is 10th December 2005 at 00:32:31 am in GMT format (Local time (India) is +5:30 GMT which is 06:02:31AM)
5. The file name is bride06dec[1].gif and the directory name where the file is stored is 0DUDEX0X
6. The HTTP headers field holds important information about the context in which the file was retrieved.
7. Using the file and directory fields, the file can be located.
8. The file bride06dec[1].gif should be available under the directory 0DUDEX0X which is present in the directory "C:/Documents and Settings/CnX/Local Settings/Temporary Internet Files/Content.IE5″. This directory has folders with random obscure names and an index.dat file which we analyzed using the Pasco utility.

Using a hex viewer such as WinHex, we can view the index.dat file and at byte offset 0×80 begins a listing of directories on the suspect's computer that contain the cache files. As we can see, 0DUDEX0X is the second directory present in the listing.

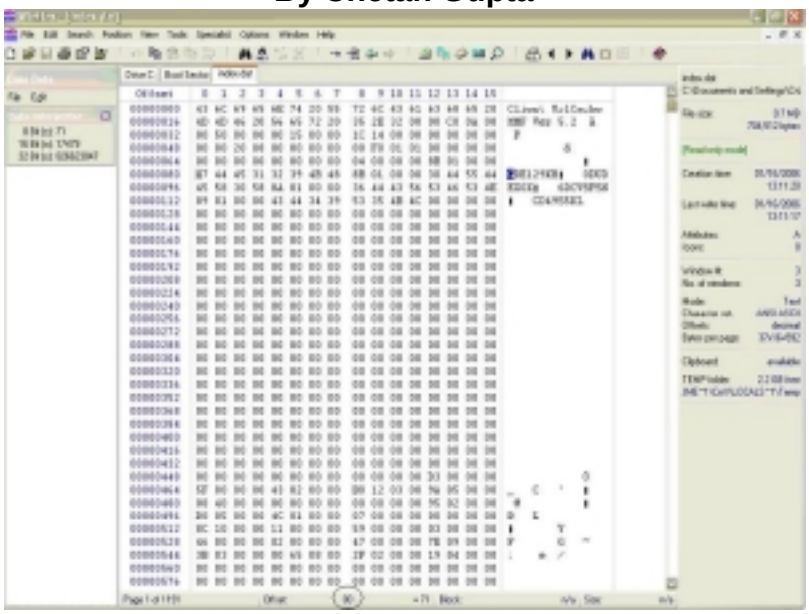

Figure 5: WinHex Output: 'index.dat' file

The same procedure can be repeated until all files for a page are located which allows us to reconstruct visisted the page completely!

Web-based Email Traces

Some of the most interesting files in the IE cache are those related to web-based email access such as Yahoo! Mail. When a user access his Yahoo! account the following files can be of interest to the investigator:

- **ShowLetter:** are actual emails being viewed in the Yahoo! inbox
- **Compose:** are cached copies of emails that were being composed
- **ShowFolder:** are cached copies of the folders in the Yahoo! mailbox

We have found vital clues in these files, especially when used with a File Recovery software. The user may have tried to clear his IE History, Cookies and Temporary Files, but File Recovery software can recover either complete files or important fragments of the cached files, as shown below:

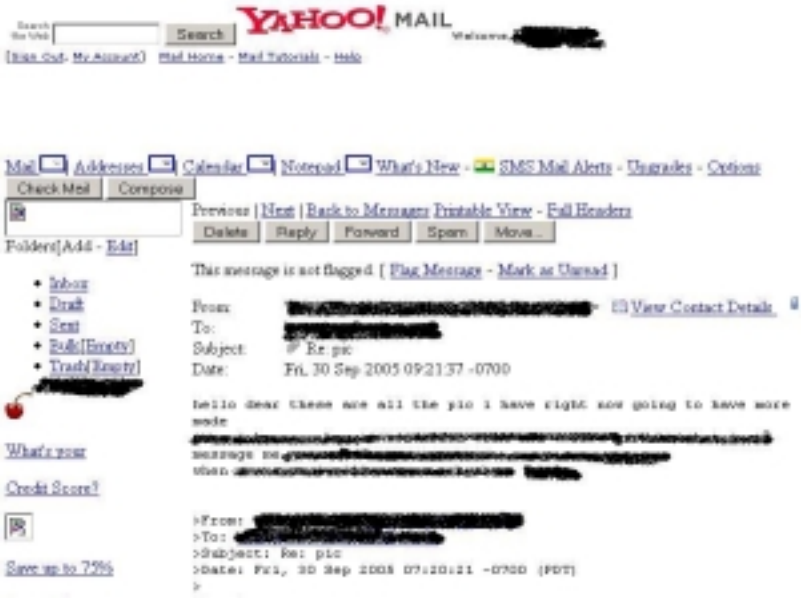

Figure 6: Yahoo cached email file

It is important to note that simply loading the ShowLetter, ShowFolder, or Compose files in your browser won't work because there is Javascript code at the start of each file, which checks if the user who is accessing this page has been logged in to Yahoo! Mail or not. If not, it redirects him to the login page. This can be easily bypassed by first opening the .html file in a text editor, and deleting any references to "login", such as the following lines:

```
<noscript>
<META HTTP-EQUIV=Refresh CONTENT="0; URL=/ym/login?nojs=1">
</noscript>
```

Investigating Cookies

The history file in the C:\Documents and Settings\profile_name\Cookies directory is identical in structure to the main index.dat files but instead of URLs, it stores the history of the cookies. Pasco can be used to parse this index.dat file. Often the mere existence of a cookie is enough to show that a user was visiting the site.

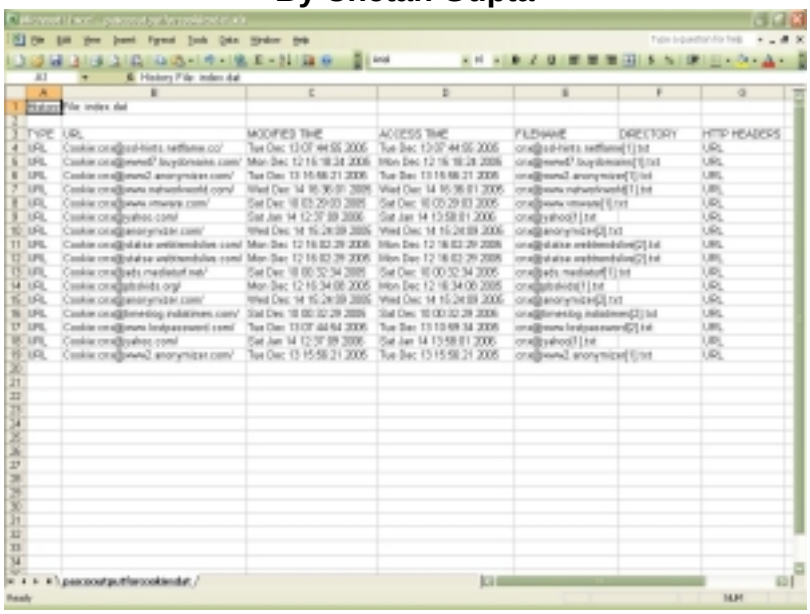

Figure 7: Investigating Cookies: Level 1

To delve deeper into the user activity and look at the information inside an individual cookie, we can use a free tool available from Foundstone called Galleta.

Tool: Galleta

Galleta is an open source IE Cookie Investigation tool developed by Keith Jones to translate the information inside a cookie.

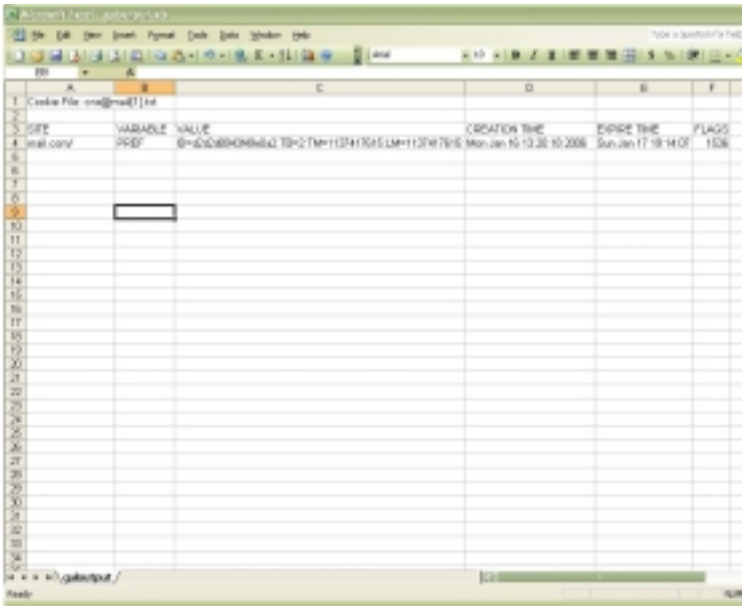

Figure 8: Investigating Cookies: Level 2

General Tools

Let us have a peek into some tools which greatly aid the investigator in reconstructing the web activities of a suspect.

BROWSWER SECRETS UNVEILED

By Chetan Gupta

Web Historian: Web Historian parses data from nearly any Web browser and comes as a native windows application. It was released by Red Cliff Consulting as a free tool and can parse cached information from a variety of browsers including IE and Mozilla.

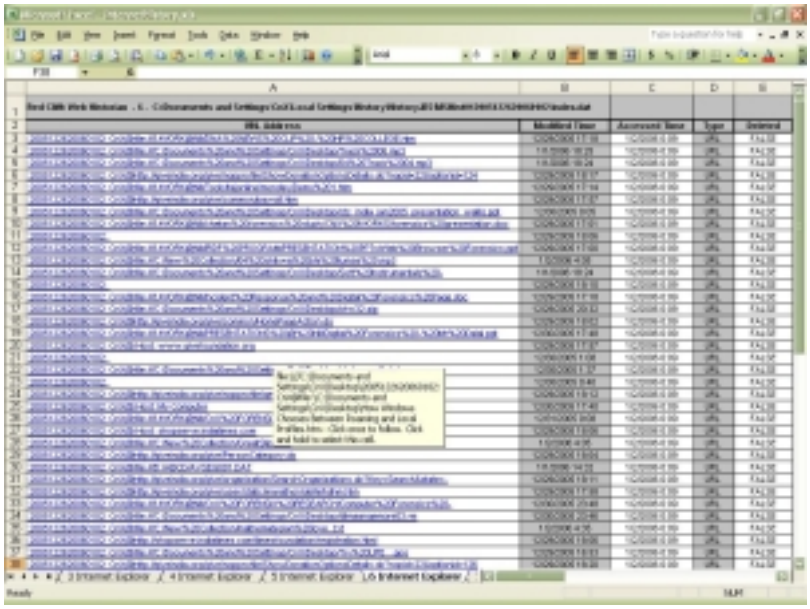

Figure 9: Webhistorian

Encase: Encase is a popular commercial tool that has a built-in functionality to examine a user's Web browsing activity. Encase utilizes a script, referred to as an E-script, to parse the Web browsing information found in the evidence and present it to the investigator.

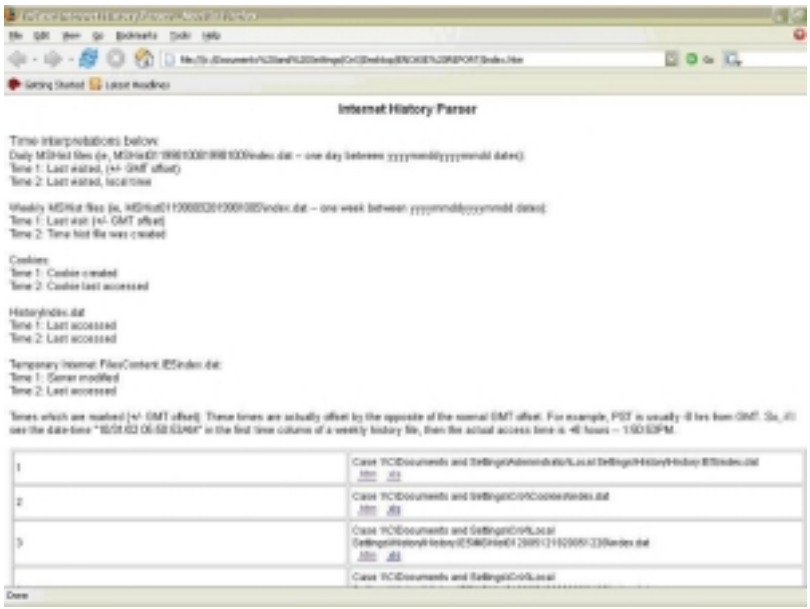

Figure 10: Encase

Afterthoughts

In this article, we have seen the various logging facilities available in the IE browser to track a suspect's web usage. Correlating time frames from all three facilities can help the investigator a great deal in identifying and understanding the suspect's activities on the web.

References

- ForensicsFocus
- Web Browser Forensics, Part 1. Rohyt Belani and Keith Jones. Securityfocus.
- RealDigitalForensics