

# PDF Forensics: Uncovering MYD files

By Ray Yepes – ISSA member, South Texas, USA chapter

**A forensic analyst discovers a procedure for uncovering and analyzing recently accessed PDF documents.**

During a recent criminal investigation, I was asked to conduct a forensic analysis on a computer which allegedly contained, or had contained, a highly classified document. More specifically, according to the federal search warrant, the investigation was to search the suspect's hard drive for any evidence of a PDF file named *Highly\_Confidential.pdf*.<sup>1</sup> The federal prosecutor alleged that this highly confidential document containing top secret information was stolen from a secured government mainframe, and that the existence of this file (because of the uniqueness of the filename), at any point in time was sufficient evidence to prosecute the offender. The defendant, of course, argued that he was never in possession of such document.

The evidence was provided in the form of a forensic image<sup>2</sup> of the offender's hard drive. One of the first steps was to obtain either a copy of the document in question in order to formulate keyword terms and/or calculate the unique hash value or a hash value of the file to run it against the evidence. Due to the sensitive nature of such document, this request was denied; all I had to go by was the name of the document in question – *Highly\_Confidential.pdf*. Considering the limitations, the next step was to proceed with an analysis and recovery of all existing and deleted PDF files while employing commonly utilized forensic software, including the recovery of PDF files from unallocated clusters. This analysis yielded no positive results.

A keyword search was configured using the document's name and possible variations (such as short name "Highly~1.pdf" or temporary instances) in case the document had been renamed by the user or changed by an application process. Normally, an investigator can expect to find numerous remnants

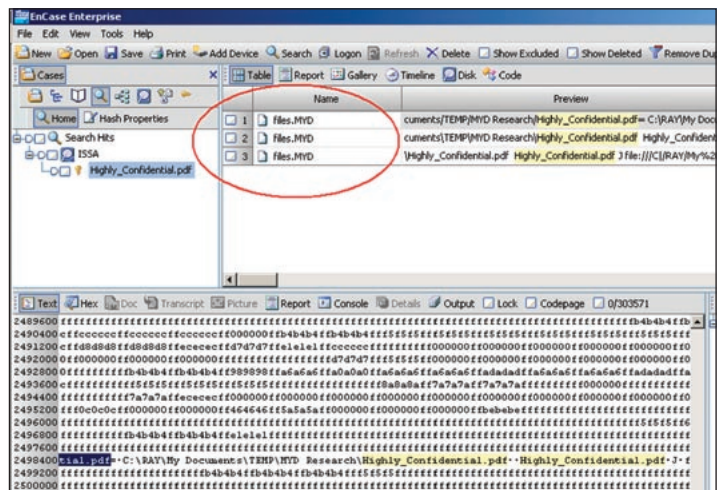


Exhibit 1

located throughout the hard drive in areas such as the pagefile, unallocated clusters, MFT records, file slack, recycle bin, registry entries, directory entries, prefetch entries of executable files, etc.<sup>3</sup> By simply opening a file, numerous evidence fragments are created and left in portions of the hard drive. Interestingly the keyword search yielded only one result as shown in Exhibit 1.

Based on forensic evidence, there was a positive hit for the filename "Highly\_Confidential.pdf" found in the following path:

```
\Documents and Settings\username\Application Data\
Adobe\Acrobat\7.0\organizer70\files.MYD
```

This result was intriguing as it is not the normal outcome one would expect, thus suggesting that perhaps an effort by the offender had been taken to conceal the records. Further analysis revealed that the offender had BCWipe<sup>4</sup> software installed, which is designed to securely delete files from disks and other media by wiping and overwriting the sectors where these files were stored, making it nearly impossible to restore

1 Actual name of document has been sanitized for the purpose of this article. In addition, all exhibits are for representational purposes only and are not actual screenshots.

2 A bit stream image copies every bit and byte on a hard drive including unused disk space and unallocated sectors.

3 Pagefile or swapfile is used for virtual memory implementation; unallocated clusters – unallocated disk sectors used in a file allocation table; MFT – Master File Table (a table of metadata); file slack – data storage space that exists from the end of the file to the end of the last cluster assigned to the file; prefetch folder is used to improve the loading time of executable files.

4 [www.jetico.com/bcwipe3.htm](http://www.jetico.com/bcwipe3.htm).

any data that has been properly wiped. Some of the functions of this software allow the user to wipe the content of the recycle bin, wipe recently used records, wipe free space, and wipe and encrypt the swap file including directory entries for FAT and MFT records. This finding would explain why only one entry resulted from the keyword search.

The next step was to figure out the purpose of *files.MYD*. The file path (See Exhibit 2) reveals it is associated with Adobe Acrobat 7.0 (not Acrobat Reader). Numerous phone calls to Adobe Systems Inc. technical support produced very little information. Further research on my part determined that *files.MYD* was some sort of database file. Examination of the Adobe Acrobat organizer feature (See Exhibit 3) showed that the file maintains a list of recently opened documents, thus substantiating it is a database file. Further research determined that the extension “.MYD” is associated with MySQL databases.

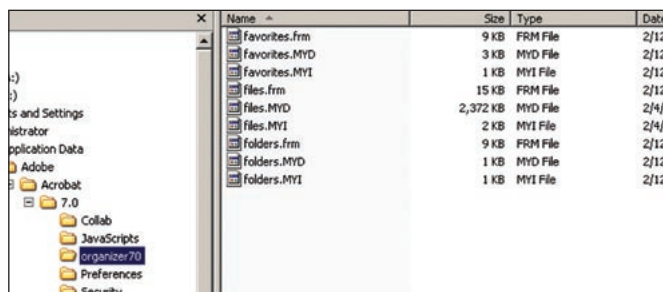


Exhibit 2

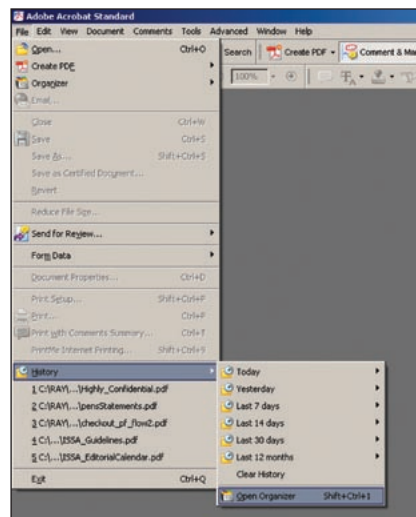


Exhibit 3

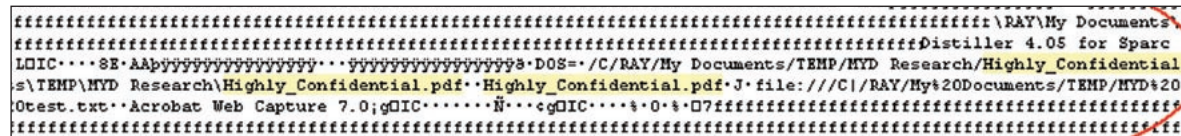


Exhibit 4

not be submissible evidence in court. Even though the prosecutor could dispute the fact that a file with a very complex filename was discovered on the defendant’s computer, no additional argument could be sustained. Consequently, I felt it

was imperative to decipher the observations. Otherwise, the finding would not have enough weight in court.

In an effort to decipher the content of *files.MYD*, I devised the following experiment. For the test environment two identical computer systems were configured with Windows XP, Service Pack 2 and Adobe® Acrobat 7.0 Standard edition. They were both loaded with a fresh computer image.<sup>6</sup> A compilation of numerous *files.MYD* files was gathered which included a file dating as far back as 2004. For the purpose of this article, a *files.MYD* file created on 02/12/06 was used. While looking at the Organizer feature of Acrobat on a live machine, I observed that records are kept for up to twelve months (See Exhibit 3). However, after further experimentation, I noticed that all records were preserved, regardless of the date stamp. I changed the date and time properties for the test machine and was able to successfully access records dating back to 2004. However, the organizer feature still displayed only 12 months worth of records based on the date of the computer system. The information provided by Acrobat Organizer was very detailed – it provided and/or retained information such as last opened date, last modified date, filename, document path, etc. – information that could prove critical for the prosecutor’s case (See Exhibit 5).

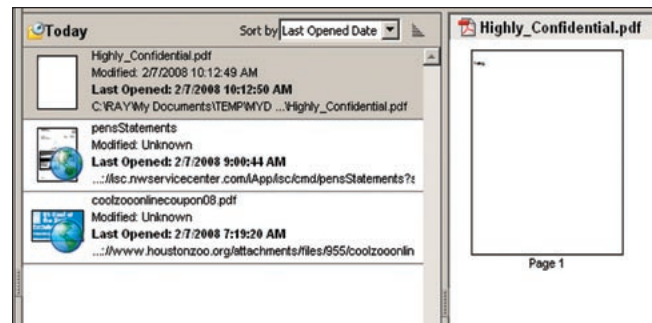


Exhibit 5

Clearly there was vital information to be found within the *files.MYD* file. It became critical to the investigation to be able to view and/or decipher the contents of this file. Revisiting the forensic tools, I exported all files located in the “organizer70” folder, including the “files.MYD” file, (reference Exhibit 2 for a listing of files located in this directory), and copied them to the “organizer70” of the other test machine. To my surprise, copying the files from one system to another did not work and the organizer displayed no files and/or recent activity

at all. Reverting the above procedure, the original files were restored back to the “organizer70” folder and then opened with Acrobat Organizer – all files and recent activity were displayed accordingly. After tweaking and trying this procedure several times, all attempts proved unsuccessful. I believe that perhaps the *files*.

5 EnCase® was used for this example.

6 Using Symantec Ghost Solution by Symantec Corporation.

MYD file maintains a unique installation ID that is verified with the product key of Acrobat, making each *files.MYD* exclusive to a specific system (this is just a theory that I have yet to validate; however, moving the *files.MYD* file or entire “organizer70” folder from one computer system to another definitely did not work during the experimentation).

Back to square one, I decided on a different approach. Knowing that the filename extension “.myd” was associated with that of a MySQL database, I downloaded MySQL<sup>7</sup> and MySQL Tools<sup>8</sup> (which includes MySQL Administrator and MySQL Query Browser). After installing MySQL in the test environment, I copied the “organizer70” (including all files) folder exported from the suspect’s computer to C:\Program Files\MySQL Server 5.0\data (See Exhibit 6).

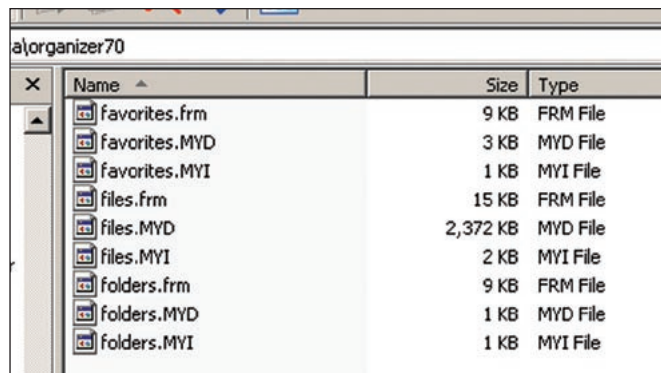


Exhibit 6

Without any modifications to the default configuration of MySQL, a server connection was configured using the MySQL Query Browser connected to the MySQL server instance (See Exhibit 7).

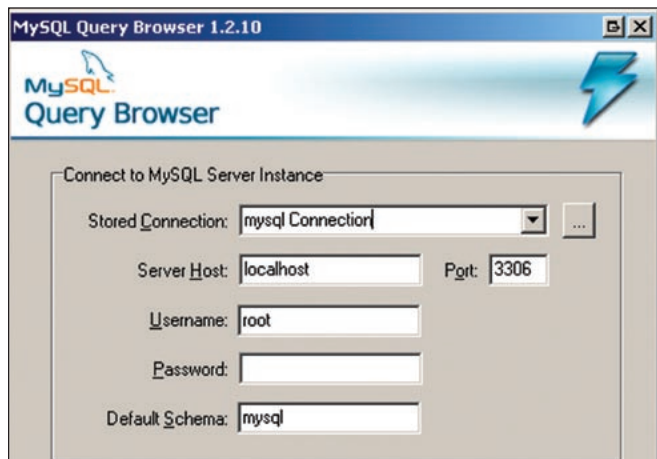


Exhibit 7

Once the connection to the server was successfully established, a link to “organizer70” folder was successful, effectively mounting the offender’s *files.MYD* file to MySQL (See Exhibit 8).

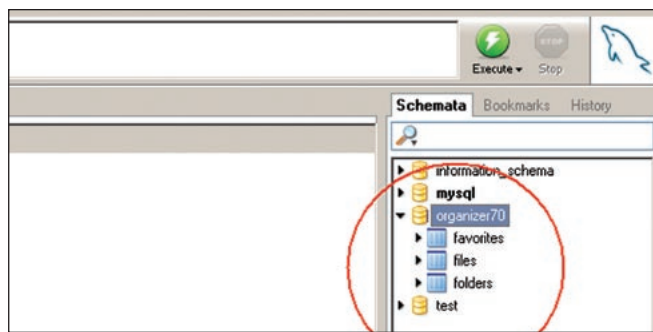


Exhibit 8

I entered SQL statement syntax `SELECT * FROM organizer70.files` in the Query Browser (See Exhibit 9).

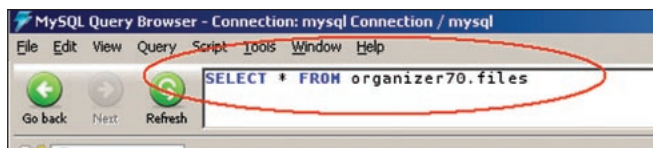


Exhibit 9

As expected, this procedure effectively decoded the *files.MYD* file content and completely accessed all records within the database (See Exhibit 10).

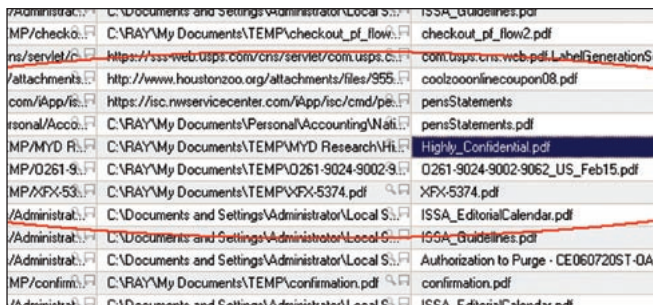


Exhibit 10

It was a total surprise to discover that the “files.MYD” database file contained additional information not displayed by Acrobat Organizer, which was limited to the last opened date, last modified date, filename, and document path. However, using MySQL Query Browser additional information was revealed, information such as author, creator, producer, keywords, display path (apparently different than document path), file system (local source versus Internet source), file size, number of pages, last viewed date, title, subject, and more. The additional information obtained from this database proved extremely useful for the prosecutor as it provided comprehensive and incontrovertible evidence that the highly classified document *Highly\_Confidential.pdf* was accessed from the defendant’s system, thus substantiating the case – the information attained was admissible evidence in court. After this evidence was presented to the court, the defendant accepted a plea bargain from the prosecutor and pleaded guilty in federal court to one count of conspiracy and defrauding the United States by accessing a government computer. Case closed.

This “files.MYD” database provides an immeasurable source of information for forensic analysts as this database maintains

7 MySQL Ver. 5.0.37 for Windows.

8 MySQL Tools Ver. 5.0.10.

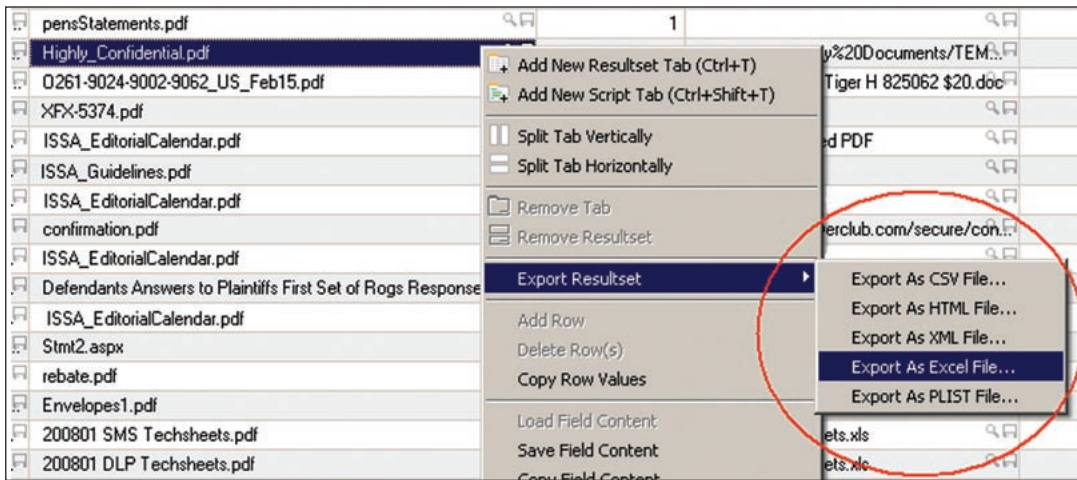


Exhibit 11

and preserves information for accessed PDF files. However, there are a few drawbacks, the first one being that this feature is only available with the full version of Adobe Acrobat products. The second disadvantage is that even though MySQL Query Browser allows a user to search for a specific filename, unless the user is familiar with SQL syntax statements, the user will not be able to sort by columns. Nevertheless, this weakness is easily overcome by the ability to export the query results to a variety of file formats such as CSV, HTML, XML, PLIST and excel file (Exhibit 11). There seems to be very little information about “.myd” files available (at least that I have been able to find); this is what prompted me to put together this article, as this technique has proven useful in many other investigations I have conducted since discovering it. Now, as routine as checking the “recent” documents folder during an investigation, I have added this procedure to my toolkit. Browsing the folder \Documents and Settings\username\Application Data\Adobe\Acrobat\productversion\ and checking for presence of an organizer folder might just prove crucial to your next investigation.

odological guide to deciphering Adobe Acrobat “files.MYD” files, and a step-by-step guide on installing and configuring MySQL server and MySQL Tools (MySQL Administrator and MySQL Query Browser), including SQL syntax statements helpful for sorting and filtering the file content. This guide is available at [www.rayyepes.com/pdffiles.htm](http://www.rayyepes.com/pdffiles.htm).

### About the Author

Ray Yepes is a security professional based in The Woodlands, Texas. With a bachelor’s degree in Computer Science from Sam Houston State University, he is currently working on his Masters of Science in Digital Forensics. He has over 18 years of experience as an IT professional and 13 years as a security analyst and criminal investigations examiner. He has assisted and advised federal, state and local government agencies in criminal investigations, high profile cases, and national security matters since 1995.



Photograph provided by Blossom Photography

Note: For the purpose of this article, all exhibits are for representational purposes only and as such were created for the intention of this article; they do not represent actual screenshots of the investigation. In addition, the more technical and procedural content has been removed. Nevertheless, I have developed a more meth-