Why Mr. Zhao didn't do it?

by Bin Xie, CISSP, CISA

1. How did the police find the child porn file on Mr. Zhao's computer?

One of the biggest questions in this case is: how did the police officer accurately and precisely find the first porn file on Mr. Zhao's computer?

According to one of the presentations of the DA's office, a "Hash" match has been used to discover the first and original evidence.

Why use hash values (such as SHA1, MD5, etc.) to search for child porn files that are stored in people's computers over the Internet? Here is why: The same porn file on the net may use different names by different people and is easily changed, so you can't use file names to search for it. Also, many files with the same format and size may contain different contents. So you can't use file size to search for the file either. How about using keyword search? It doesn't work. The graphic and video files are not text files!

Fortunately, in this world, each electronic file has its own unique secret value, and we call it "hash value". Even if you change one byte of an electronic file, the file's hash value will change greatly. As long as you don't modify the file, multiple copies of the same file have the same hash value. This can help people accurately find the exact files on the net.

So if a police officer gets a copy of a child porn file online, the officer will take its hash value, then put it into a child porn database. After that, a police computer is used to search all files on the net, and find files on the net that have exactly the same hash value. You may end up finding multiple copies of the same child porn file on the net if your computer is powerful enough to conduct a large-scale search.

It is easier said than done. How many files are on the Internet? It is not counted by billions, trillions, it is counted by …… who knows!

Here are 3 ways to improve the search:

(1) Create a new type of search engine to search and find all files on the net, and then assign each file a hash value. Now you just need to let it find the files of those have a hash value that matches the porn file's hash value in the database. Unfortunately, this search engine has not been made yet.

(2) Ask all file owners/publishers in the world to attach hash value to each file. So you can use an existing search engine to do the search. Unfortunately, no one will do it.

(3) Make a dedicated search tool to search only for a few types of files (such as .avi, .jpg, .mp4, .gif), and for each file it finds, check its hash value. If any file's hash value matches one of the 4 million child porn hash stored in law enforcement's database, take a note of the IP, and then take actions. This is likely what the law enforcement is using.

However, there are many Limitation of such dedicated search software.

- It only works in small scale like a corporate network due to the limited capability.

- Even local area's (like a city or a state) law enforcement's efforts may easily go beyond their boundary. A data center located in one city may serve customers living in other cities or states. Therefore such search engine must remain small scale.

- A search covering billions of trillions of files on the net against 4 million hash values in its database will take tremendous efforts because, for each file it finds, it will calculate its hash value and then save it into its database. Next, compare it to the 4 millions of child porn to see if there is a match or not. Therefore this broad search is very, very ineffective and slow. It is like checking millions of different foods against tens thousands of chemicals they contain.

2. How could police officer locate the child porn file on Mr. Zhao's computer so fast?

According to the police forensic report, the initial child porn file was somehow got into Mr. Zhao's computer on 1/3/2015. Two days later, on

1/5/2015, a police officer successfully downloaded the file from Mr. Zhao's computer via eMule or eDonkey file sharing. As we discussed previously, the officer uses SHA1 (a type of hash) to search and find the file. This could be true if it takes 2 months. The fact is: It is impossible to conduct a large-scale search and find it in 2 days!

Another possibility is that the police officer may just installed an eMule client or eMule emulator software and started using its search feature to find the porn files on the net, and got it by accident or by luck. However, the remote end – eMule clients, don't return a hash value to another party – it returns with file name match! Therefore, the officer must have used file name search instead. As we discussed previously, this is not how the child porn file was located. This can be used to download evidence, but not to discover and find the evidence.

The biggest problem with file name search to find the first porn file on Mr. Zhao's computer is: the file name is in Chinese! These are the questions:

- Does the police officer understand Chinese?
- If he doesn't know the Chinese language, how could he use the Chinese language to search?
- Does his computer have a Chinese input software?
- Does he know how to use it? Does he know the difference between simplified Chinese and traditional Chinese?
- Does he know the file name is in simplified Chinese or traditional Chinese?

So this rules out the possibility the police officer used file name search to discover the porn file on My. Zhao's computer.

According to the logs provided by the Police, here is how they conducted the search and discovery, and found the file:

A police officer uses a software to emulate e-Mule. This software connects to a remote e-Mule client and starts to download a particular porn file in the way of 1-on-1 download (avoid connection with other e-Mule clients while downloading). After a successful download, it calculates the hash value. According to theory, it is supposed to compare it against the 4 million child porn hash values in the database and find a

match. Instead, it checks the hash value of the downloaded file against ONLY ONE hash value. This hash value is called "Expected hash". In this case, it is 3605BA1A6254E3BF4CE9741D72ADEFF7.

According to the logs provided by the Police, here is how the first child porn evidence was obtained:

On 12/23/2014， 2 unsuccessful searches were conducted against Mr. Zhao's computer to find a file with expected hash=F99AA599CFA1803C267145E297C3DDB3

On 12/25/2014， 9 unsuccessful searches were conducted against Mr. Zhao's computer to find a file with Expected hash=F99AA599CFA1803C267145E297C3DDB3, FB1A8596EA781253A205B693452BDB56, 76459EF748CFD71DC27835E868C86511, 115E2DB7DB5BCE7D080324A666795DA5, etc.

On 12/26/2014， 7 unsuccessful searches were conducted against Mr. Zhao's computer to find a file with expected hash= 4EE624CACFC3DF4A290FCF2AFB591289, 115E2DB7DB5BCE7D080324A666795DA5, F99AA599CFA1803C267145E297C3DDB3, 76459EF748CFD71DC27835E868C86511, 115E2DB7DB5BCE7D080324A666795DA5, etc.

On 12/27/2014， 3 unsuccessful searches were conducted against Mr. Zhao's computer to find a file with expected hash= 76459EF748CFD71DC27835E868C86511, 4EE624CACFC3DF4A290FCF2AFB591289, and 115E2DB7DB5BCE7D080324A666795DA5.

On 12/28/2014， 2 unsuccessful searches were conducted against Mr. Zhao's computer to find a file with expected hash= 4EE624CACFC3DF4A290FCF2AFB591289

On 12/29/2014， 14 unsuccessful searches were conducted against Mr. Zhao's computer to find a file with expected hash= FB1A8596EA781253A205B693452BDB56，

4EE624CACFC3DF4A290FCF2AFB591289，
F99AA599CFA1803C267145E297C3DDB3，
76459EF748CFD71DC27835E868C86511, etc.

On 12/31/2014， 3 unsuccessful searches were conducted against Mr. Zhao's computer to find a file with expected hash= F99AA599CFA1803C267145E297C3DDB3.

On 1/5/2015， 3 unsuccessful searches were conducted against Mr. Zhao's computer to find a file with expected hash= D4685E3A42A3FEF5A431D15EEA8F0E96, and ONE successful search was conducted against Mr. Zhao's computer to find a file with expected hash **3605BA1A6254E3BF4CE9741D72ADEFF7**. This is the evidence the police officer used to get the search warrant!

(Please note that I use the same color for the same hash value)

Here is the summary log:

2015-01-05 19:54:57 -- Log file created at 2015-01-06 00:54:57 Coordinated Universal Time (UTC)
2015-01-05 19:54:57 -- RoundUp eMule version 1.52
2015-01-05 19:54:57 -- Started single source download thread (94B64168801001)
2015-01-05 19:54:57 -- Remote client address is: 65.78.83.141:56663
2015-01-05 19:54:57 -- Requested name for local file: 3605BA1A6254E3BF4CE9741D72ADEFF7
2015-01-05 19:54:57 -- Expected hash=3605BA1A6254E3BF4CE9741D72ADEFF7, expected size=20301246
2015-01-05 19:54:57 -- File is calculated to have 3 parts based on expected size
2015-01-05 19:54:57 -- Remote client has responded that it has a file with eD2k hash 3605BA1A6254E3BF4CE9741D72ADEFF7
2015-01-05 19:54:57 -- Remote client uses the following file name: 苏联惹小女孩（3分钟） (1)大乳 奶 精品 收藏 无码 穴 女 少妇 强奸 另类 后门 暴 迷药.avi
2015-01-05 19:54:57 -- Remote client indicated it has the entire file; assuming 3 parts based on expected file size of 20301246 bytes

2015-01-05 19:54:58 -- Remote client hash set response confirms file eD2k hash: 3605BA1A6254E3BF4CE9741D72ADEFF7
2015-01-05 19:54:58 -- Remote client hash set response confirms the file size: 20301246 bytes
2015-01-05 19:54:58 -- Remote client hash set responded with the expected number of file part hashes: 3
2015-01-05 19:54:58 -- Name given to local copy of file: 3605BA1A6254E3BF4CE9741D72ADEFF7.avi
2015-01-05 19:54:58 -- No longer connected to the remote client
2015-01-05 20:06:34 -- No longer connected to the remote client
2015-01-05 20:32:28 -- No longer connected to the remote client
2015-01-05 21:01:51 -- No longer connected to the remote client
2015-01-05 21:01:51 -- Downloaded 3 parts: remote client possessed 3 parts; file total is 3 parts
2015-01-05 21:01:51 -- Downloaded file has an eD2k MD4 hash of 3605BA1A6254E3BF4CE9741D72ADEFF7
2015-01-05 21:01:51 -- Hash value matches expected value
2015-01-05 21:01:51 -- SUCCESS: downloaded file has the expected eD2K hash
2015-01-05 21:01:51 -- Other hashes calculated for downloaded file: MD5=1702470AEF65ECF8D092ABA6697D1E4B, SHA1=F7CD9E684DB1BFA38B31C37B0BA912649DFB5FE3
2015-01-05 21:01:51 -- Thread is terminating (94B64168801001)

By analyzing the logs, here are some findings:

(1) The police were NOT conducting the random search. Instead, they were conducting targeted searches – several hash files were searched multiple times. These logs suggest that it is very likely they already knew what's on Mr. Zhao's computer even before the search and discovery started.

(2) The police were able to know ahead of time that file **3605BA1A6254E3BF4CE9741D72ADEFF7** already exist on Mr. Zhao's computer. They didn't search for it before 1/5/2015, they searched for it after 1/3/2015. Why? One obvious answer is that they knew the file landed on Mr. Zhao's computer on 1/3/2015. Therefore they didn't bother to search for it before 1/3/2015! The biggest question is: how did they know this ahead of the time? As we concluded previously, very likely, the police officer doesn't

understand Chinese, but the file name is in Chinese! How did he find out?

(3) The fact is, Mr. Zhao did NOT know anything about 3605BA1A6254E3BF4CE9741D72ADEFF7. I will prove it in the next section.

3. Other key information that has not brought attention to all parties or been purposely hidden by the police

There is overwhelming information that has not been brought to the Court during the trial, such as:

(1) The operating system of Mr. Zhao's computer is Windows XP, which is no longer supported by Microsoft and is extremely vulnerable to attacks, hacking, remote intrusion, unauthorized remote manipulation, etc. Anyone with some basic hacking skills and some free tools can easily take control of the computer without the owner's authorization.

(2) There is no user login name/password to access the computer. It is very likely that the built-in accounts' (such as "Administrator") passwords are all blank.

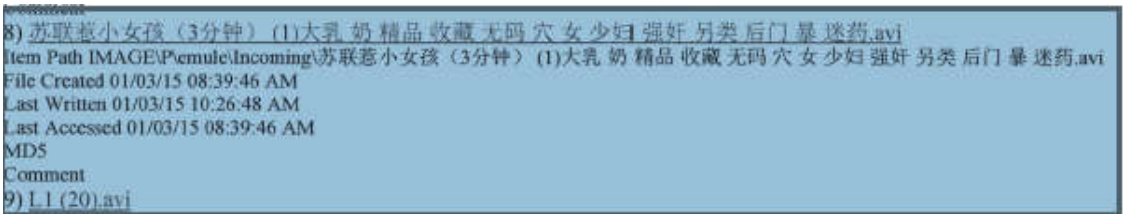
(3) On the forensic report, there is no information about any of the local logs. Why? Without the logs, it is difficult to know exactly what happened.

The biggest puzzle regarding the evidence in this case is: Internet access activity logs! The police obtained a court order to find Mr. Zhao's home address by the IP address of the source of the child porn file. With the same court order, the law enforcement can easily obtain the network access activity logs of Mr. Zhao. With such information, the police officer can convince the Court where Mr. Zhao got the file from. That evidence will make the case much stronger. Let's don't forget: the first discovered porn file on Mr. Zhao's computer was somehow got there on 1/3/2015, the ISP used the IP to track back to Mr. Zhao's address after receiving court order. There are two types of logs Internet Service Providers normally keep. The first type of logs is the IP assigned to each customer's router or modem when the router/modem is turned on/off. This type of logs can be kept for

up to 2 years because it doesn't take much space. TIME WARNER CABLE stores IP-address logs for up to 6 months. Verizon's IP address assignments are retained for 18 months, which is one of the longest. What we are talking about here is the customers' Internet access logs - most Internet service providers keep the logs for business reasons (marketing). This kind of logs are usually kept only for a few weeks, at most 2 or 3 months due to the huge storage space it takes. In Mr. Zhao's case, the police can and should obtain the access history logs if they really want to prove their case. Somehow they didn't do it! Now it is impossible to obtain such logs from the ISP (in this case RCN) after three years. Why the police officers purposely ignored this extremely valuable evidence? Also, why they did not serve search warrant for the serious crimes until THREE MONTH LATER(April 21, 2015)?

One obvious answer is that the police officers intentionally wanted the evidence for planting on Mr. Zhao's computer to disappear! As a matter of fact, the police officers successfully obtained the first evidence on January 5, 2015. A search warrant or court order for child porn case normally can be obtained on the same day. The ISP (RCN) should have no difficulty to comply with the court order and surrender their Internet access activity logs to the government. The biggest problem is: The police officers did not serve search warrant until THREE MONTH LATER (April 21, 2015). Why? Obviously they wanted to wait until related activity logs to be deleted by RCN, so that anybody including Mr. Zhao's attorney was unable to obtain the related activity logs.

Another big puzzle in this case is: According to the forensic report, since the moment the porn file landed on Mr. Zhao's computer, no one has ever, ever accessed it except the police officer! Why? Because the File Created time and Last Accessed time are the same! If Mr. Zhao had downloaded the porn file on his computer, why he has never opened it? If he doesn't even view it, then what's the point of downloading it? The only possibility is that he didn't download the file at all, and doesn't know anything about it at all! Someone else did it!


Comment
8) 苏联惹小女孩（3分钟）(1)大乳 奶 精品 收藏 无码 穴 女 少妇 强奸 另类 后门 暴 迷药.avi
Item Path IMAGE\P\emule\Incoming\苏联惹小女孩（3分钟）(1)大乳 奶 精品 收藏 无码 穴 女 少妇 强奸 另类 后门 暴 迷药.avi
File Created 01/03/15 08:39:46 AM
Last Written 01/03/15 10:26:48 AM
Last Accessed 01/03/15 08:39:46 AM
MD5
Comment
9) LI (20).avi

4. Summary of Findings

- The initial evidence cannot be found that fast (within 48 hours) by a broad hash search using a search engine.

- The police started searching Mr. Zhao's computer for child porn files on 12/23/2014. The searches were targeted at specific child porn files on Mr. Zhao's computer.

- Using file name search is impossible to find the file because the officer does not understand Chinese.

- The police already knew something about the porn file on Mr. Zhao's computer EVEN before the search started. Otherwise, the police wouldn't conduct such accurate and precise searches. The indication is: the policy already knew what's on Mr. Zhao's computer. But how? If Mr. Zhao shared the porn file to the public, the search is legit. For the files Mr. Zhao didn't share, the search must follow the law - either with permission from Mr. Zhao or with a court order. Neither was obtained.

- The computer has no password, and the operating system is extremely vulnerable.

- The forensic report doesn't provide local log analyzing results.

- The police officer has no interest to obtain activity logs from either the local computer or from the ISP which can show how the file was downloaded to that computer. Why? Can we say that the police officers intentionally let the key evidence for planting the file on Mr. Zhao's computer to go away?

- Finally, the porn file was not even opened after download, never! Mr. Zhao has no information about the porn file!

All above information points at one possibility: These evidence were planted, Mr. Zhao didn't do it, someone else did it! Mr. Zhao is framed.

Then who did it? To find out who did it, the network activity logs must be obtained! Why the police had no interest to obtain such important evidence when they had plenty of time?

If the police doesn't have interest to obtain the Internet access activity logs, this is like suddenly finding a dead body in someone house that came from somewhere else, and the owner or primary resident of the house has no clue about the dead body even exist, and the police has no interest to figure out how the dead body got in there the first place. Is this strange?

With all the information above, that makes the police prime suspect!