FEATURE 2

THE FALSE POSITIVE DISASTER: ANTI-VIRUS VS WINRAR & CO

Andreas Marx
AV-Test.org, Germany

In October 2003 I wrote an article for *Virus Bulletin* about false positives in anti-virus software (see *VB*, October 2003, p.17). To be more exact, the article was about viruses being reported by scanner A in the program or data files of scanner B – and *vice versa*. This problem was caused mainly by unencoded virus scan strings and disinfection routines (e.g. registry keys and files which should be removed) in addition to overzealous heuristics.

COLLECTING FILES

Two years later, we have built up a collection of more than 15,000 GB (15 TB) of clean files in order to enhance our false positive tests. We used two main sources for these files: first, we read in some 10,000 CDs and stored a copy of the ISO images on several storage systems. Secondly, we are mirroring more than 150 different FTP servers and downloading all new files on a daily or weekly basis.

Having such a huge test set creates some problems. For example, a couple of well-known companies have indeed released viruses or other malware together with their software. However, the number of such files is small (about 150) and insignificant compared to the several billion clean files. We left the infected files inside the collection, as all virus scanners should flag them – and if they don't, we know that some scanner tasks might have failed. A couple of these files seem to have been infected by CIH in the past and subsequently cleaned, without removing all parts of the virus, which disqualifies them for both true and false positive tests.

One of the bigger problems is related to the fact that several AV companies release updates at least once a day or even on an hourly basis. This means that test results become outdated rather quickly, since the PCs used for such a test – 15x Pentium IV 2.8 GHz and 15x Athlon 64 3500+ – would require a couple of days to scan the whole collection, for just one scanner. If it took an average of one week per scanner (taking into account common problems like crashes and required restarts) we would need more than half a year just to test the number of scanners that are included in *Virus Bulletin*'s latest VB 100% tests.

TROUBLE WITH WINRAR

To get around this problem, we focused on some key areas only. In the past I have had a couple of discussions with the author of *WinRar*, in particular about enhancing the virus protection in *WinRar* (some malware uses RAR archives instead of just ZIP files) and about a lot of false positives in his software, caused by anti-virus software. The latest *WinRar 3.50* readme file reads:

> '[…] 7. SFX modules: a) SFX modules are not compressed by UPX anymore, so they are larger now. UPX compression caused numerous false alerts by antivirus software. If you wish to use compressed modules, you can get UPX from http://upx.sourceforge.net and compress *.sfx files in WinRAR folder […]'

This was the first interesting test item where all of the scanners could be covered: a scan of the files from ftp://ftp.rarlab.com, which we had been monitoring for a couple of years. So we should have a copy of almost every version of *WinRar* released. We limited our scans to 896 EXE files (877 MB) inside the '/rar' subdirectory of the FTP server where copies of *PocketRar*, *WinRar* and some additional software can be found – and we were a little shocked by the results.

THE FIRST TEST-RUN ...

On 21 August 2005 we tested AV tools from the following vendors:

AntiVir (*H+BEDV*)	*Kaspersky*
Avast (*Alwil*)	*McAfee*
AVG (*Grisoft*)	*NOD32* (*Eset*)
BitDefender (*SOFTWIN*)	*Norman*
ClamAV	*Panda*
Command (*Authentium*)	*Proland*
Dr.Web	*Proventia-VPS* (*ISS*)
eSafe (*Aladdin*)	*QuickHeal*
Fortinet	*Sophos*
F-Prot (*Frisk*)	*Symantec*
F-Secure	*Trend Micro*
Hauri	*VirusBuster*
Ikarus	*eTrust-INO* & *eTrust-VET*(*CA*)

Of the 27 scanners tested, six reported up to 111 infections and two of them reported up to 709 'suspicious' files (see Table 1, left-hand column). Some examples:

- *Avast* reported that a 'sign of "Win32:Trojan-gen. {UPX!}"' was found in the file 'wrar300r.exe\Zip.sfx'.
- *AVG* found 'wr330sc.exe – Trojan horse Agent.M'.
- *ClamAV* reported – 'wr341ro.exe: Oversized.RAR FOUND' and 'wr32b1el.exe: Trojan.Spy.Banker.C*Y FOUND*'.

Program	Infected [suspected] files 1st scan 2005-08-21	Infected [suspected] files 2nd scan 2005-09-11
AntiVir	0	0
Avast	10	0
AVG	1	0
BitDefender	0	0
ClamAV	111	111
Command	0	0
Dr.Web	0	0
eSafe	0 [203]	0 [203]
eTrust-INO	0	0
eTrust-VET	0	0
Fortinet	10 [709]	0 [244]
F-Prot	0	0
F-Secure	0	0
Hauri	0	0
Ikarus	1	1
Kaspersky	0	0
McAfee	0	0
Nod32	0	0
Norman	0	0
Panda	0	0
Proland	0	0
Proventia-VPS	0	0
QuickHeal	9	8
Sophos	0	0
Symantec	0	0
Trend Micro	0	0
VirusBuster	0	0

Table 1: False positives caused by the different AV tools in case of files from ftp://ftp.rarlab.com/rar/.

- *eSafe* complained about the file 'wr341cz.exe – Infected with suspicious Trojan/Worm'.
- *Fortinet*'s detection included '"wr341cz.exe" is infected with the "W32/PoeBot.D-bdr" virus' and '"wr34b1tr.exe" is infected with the "W32/Bancos.GP-tr" virus'.
- *Ikarus* reported 'wr311sc.exe – Signature "Win32.Elkern.C" found'.
- One of *QuickHeal*'s findings was 'pk33b1.exe – Infected : (TrojanSpy.Bancos.B)'.

From an initial look at the 'malware names', it appears that (self-extracting) *WinRar* archives are often used for the packaging of malware, like password-stealing Trojans or Backdoor programs. Furthermore, it looks like some signatures are simply not created properly, which causes false positives when the *WinRar* stub is found. Besides this, some scanners are a little over-paranoid with their heuristics and create too many false positives if files are simply runtime-packed.

The scan time was very interesting too, as some tools were really checking all files inside the *WinRar* archives (which were self-extracting RAR files most of the time), while others only checked the small Win32 stub of the SFX archive, without inspecting the files inside.

For example, *Sophos* proved to be the fastest scanner with a scan time of only 30 seconds, *Trend* took 40 seconds, *Fortinet* about 700 seconds (11.5 minutes), *BitDefender* around 750 seconds (12.5 minutes), *Kaspersky* 1,300 seconds (22 minutes), *Hauri* about 2,400 seconds (40 minutes) and *Proventia-VPS* 3,200 seconds (53 minutes). It should be noted that *Proventia-VPS* is not a virus scanner working with signatures, but a behaviour-based product which requires a longer scan time. From the scan time requirement, one can easily see which of the scanners really inspected all 42,843 files inside the 896 (self-extracting) EXE files. If an AV program doesn't scan the whole self-extracting *WinRar* archive, it is not able to find infected files inside it and thus, it's also less likely that false positives are caused.

... AND THE SECOND TRY

On the day of our initial test, we notified the AV companies, discussed the results with them and provided samples of the files to those who requested them. Then, on 11 September – exactly three weeks after the first test – we repeated the false positive test with the same set of files (no new *WinRar* versions had been released in the meantime).

The number of trouble-makers had decreased significantly, but there were still a lot of files flagged as being 'infected' or 'suspicious' by many of the tested programs (see Table 1, right-hand column).

All of these AV companies were notified again, of course. The high number of false positives generated by *ClamAV* can certainly be considered critical. However, the 203 'suspicious' warnings by *eSafe* and the 244 which were left by *Fortinet* are not really good either.

THE COST OF FALSE POSITIVES

It seems to me that files need to be processed more carefully, especially in the case of installers (like the *WinRar* stub) or runtime engines, as illustrated by the following example.

A well-known computer magazine contacted me on 30 August regarding the games ActionBall 2003, ActionBall 2004 and Jumpy Balla 2003, which can be found at http://www.happy-future-software.de/. Of the 27 AV tools

tested, six found a virus inside the files. *Dr.Web* told us it had found 'Win32.HLLW.Franvir', *F-Secure* and *Kaspersky* both complained about 'P2P-Worm.Win32.Franvir', *Ikarus* found 'P2P-Worm.Win32.Franvir', *NOD32* showed an infection of the 'Win32/Franvir.C worm' and *VirusBuster* reported 'Worm.P2P.Franvir.B'.

After issuing our report, we received a response from *F-Secure*, explaining that the file had been created by GameMaker 4.3, which was also used by the Franvir worm. An email from *Kaspersky Lab* arrived just a few minutes later, explaining that all games created by this tool use the same interpreter stub – it is just the data segment with the game logic that differs. An email from the *NOD32* team arrived three hours later, confirming the false alert and indicating that it will be fixed with the next engine update. However, none of the other companies responded or fixed it.

This false positive was rather significant, as the computer magazine had just produced several hundred thousand cover CDs which included these games. Just a couple of hours later, they would have destroyed all the CDs to make sure they were not about to distribute possibly infected software. With the resulting delay in shipping the magazines (it would have been necessary to remove all 'old' CDs manually and newly created CDs would have had to have been stuck in) and the cost of creating new CDs, the magazine estimated that the damage caused by the false positive could easily have reached a level of several hundred thousand euros.

CONCLUSION

A lot of AV companies have automated the process of creating signatures for static malware. Due to the fact that a lot of malware uses *WinRar* self-extraction archives at some point, the number of false positives had been growing rapidly in this area. False positives could not only prove costly for companies if they find some 'suspicious' tools on their hard disk, but the case of the magazine cover CDs illustrates how else false positives can have a significant impact on businesses. It should be noted that *WinRar* and *GameMaker* were just two examples of what could be many more.

Therefore, a large collection of 'known good' files is essential in order to create high-quality software. Some of the smaller commercial AV companies and the developers of the Open Source project *ClamAV* urgently need to do something in this area.

While well-working processes already exist in order to report new malware and add detection for those files, it is important to attain the same high quality of processes in the case of false positives. This will hopefully reduce the impact of false positives in future and we will be able to remove files causing false alarms faster than ever before.