RETROSPECTIVE TESTING - HOW GOOD HEURISTICS REALLY WORK

Andreas Marx
GEGA IT-Solutions GbR, Klewitzstr. 7, 39112 Magdeburg, Germany
Tel +49 391 6075466 • Fax +49 391 6075469 • Email amarx@gega-it.de

ABSTRACT

Currently, there are no exact test details available as to how well the heuristics of a virus scanner really work. A number of marketing people still claim that their program's heuristics can detect up to 95% of all new viruses. However, a number of existing small-scale tests by Joe Wells for PC World *[1], the University of Hamburg [2,3], and the University of Magdeburg for the German* c't *magazine [4,5] demonstrate that this cannot be the truth. A more realistic value should be something between 15% and 55% for most scanners – and this explains the need for at least weekly updates.*

Currently, there is no exact retrospective test available, but we (AV-Test.org) want to fill this gap now. In order to do this, we have collected all available anti-virus updates of about 20 programs over a period of more than one year for the program, engine and signature updates (currently about 100 GB of compressed Image files).

The results of both an ItW test – based on all monthly published WildLists – and a few more zoo tests show how rapidly an anti-virus program becomes outdated and the development of heuristics in the past up to today on a 'general' and a 'per product'-basis. The use of the word 'heuristic' here should not only include dynamic or static methods to detect new viruses, but also generic, less exact detection of possible variants. And even random findings in some cases.

During the pre-tests for this paper, which are published in [5], with three-month old scanners against new Zoo viruses we found that macro virus heuristics – or at least a generic detection – are very well developed in most programs: 74% to 94% detection rate with an average of 86% for all programs. More room for improvements is related to script viruses, but the detection scores are still quite ok: 35% to 82% with

an average of 58%. Some programs shows very good results for Win32 file viruses, others find just a few of the new ones: 24% to 79% with an average of 57%. And finally, Win32 worms, backdoors and Trojan horses causes still the most significant problems in a lot of different scan engines: 8% to 37% with an average of 19%. Mainly the complex structure of these often High Level Language (HLL)-based malicious code causes this trouble – a deep scan of the included functions would simply be too time-consuming.

The test results with a six-month old scanner against new Zoo viruses are sometimes significantly worse, in other parts we saw just a negative change of a few percent. For macro viruses the worst program detected only 47% of all viruses while the best one was still able to detect about 89% of the new macro malware. The average was 75% detection for macro viruses. For script viruses, the detection was from 17% to 74% with an average of 43%. For Win32 file viruses the range was from 8% to 68% with an average of 37%. And again, the Trojan horse and backdoor detection score was still only about 3% to 26% with an average of 12%.

The paper will not only discuss the results of the bigger test, but also the limitations of such tests (what does it really show?) as well as a few conclusions. However, it is still 'work-in-progress' and was not finished at the deadline for these conference proceedings. The complete paper, all XLS sheets for the different programs as well as the presentation can be found at http://www.av-test.org/ during the Conference.

REFERENCES

[1] Sean Captain, 'Stealth Fighters', Online version of the September 2001 issue of *PC World,* http://www.pcworld.com/reviews/article/0,aid,55803,pg,1,00.asp.

[2] Klaus Brunnstein *et al.,* 'Heureka Anti-Virus Test July 2001', http://agn-www.informatik.uni-hamburg.de/vtc/en0107.htm.

[3] Klaus Brunnstein *et al.,* 'Heureka-2 Anti-Virus Test March 2002', http://agn-www.informatik.uni-hamburg.de/vtc/en0203.htm.

[4] Andreas Marx *et al.,* 'Comparison Test 2000-11 (Clients), Windows 98, ME - c't', http://www.av-test.org/sites/test_all.php3?test=2000-11&lang=en. Published in: Andreas Marx, Patrick Brauch, 'Schädlingssuche - 14 Antiviren-Programme im Vergleich', *c't* 02/2001, p.102.

[5] Andreas Marx *et al.,* 'Comparison Test 2002-05 (Clients), Windows ME, 2000, XP Professional - c't', http://www.av-test.org/sites/test_all.php3?test=2002-05&lang=en. Published in: Axel Vahldieck, Andreas Marx, Patrick Brauch, 'Wer suchet, der findet - 12 aktuelle Virenscanner für Windows im Test', *c't* 13/2002, p.176. Available online at: http://www.heise.de/ct/02/13/176/.