Anti-Virus Outbreak Response Testing and Impact

Andreas Marx
AV-Test GmbH

http://www.av-test.org
# Table of content

- Introduction: past and current tests
- About the project (how it works)
- Current problems
- Some interesting general statistics
- Results of the heuristic test (based on a retrospective approach)
- Results of the outbreak response time test ("who performed best" in the past eight months)
Introduction: past and current tests (I)

- Current tests are still too focused on detection rates of viruses (regardless if they are ItW or Zoo viruses)
- One addition: retrospective tests (old scanners were tested against the most current malware to see how many of them were detected proactively, without the requirement of updates)
Introduction: past and current tests (II)

- New test strategy for today’s problems: Outbreak response time tests
- The main question we want to answer:
  How long does it take until signature updates (from the different AV companies) are publicly available (using recommended downloads) in cases of major worm outbreaks?
About the project (how it works)

- Project started in the current state at 2004-01-01 (the first beta implementation was running since 2003-10-27), see VB 02/2004
- We monitor 24 different AV companies for the release of new regular and beta signatures, engine and program updates
- Checks are performed every minute since 2004-06-29 (formerly, they were performed every five minutes)
About the project: download process

- Download system is running on Debian Linux 3.0
- We only download new (changed) files, using wget
- All files are stored into a large archive in our lab (sync’ed on-demand using rsync over SSH)
- A PostgreSQL database entry is created for every download with information about the filename, the size, MD5, plus the date/time of the download
- The system is located in a data center of a big ISP with a direct 100 MBit Internet connection
About the project: test process

- First idea: manual checks in case of outbreaks (we expected only a few per year…)
- After the Mydoom / Bagle / Netsky war was started, we switched to automated tests using command-line scanners whenever possible (with the same settings the GUI version uses, e.g. for heuristics)
- We have several scan systems which are running on Windows XP and Windows 2003 Server with Cygwin (due to the use of Unix shell scripts)
About the project: participants

- Regular definition updates from AntiVir, Avast, AVG, Bitdefender, ClamAV (since 2004-02-20), Command, Dr. Web, Esafe, eTrust (CA), eTrust (VET), Fortinet (since 2004-04-15), F-Prot, F-Secure, Ikarus, Kaspersky, McAfee, Norman, Panda, Quickheal, RAV, Sophos, Symantec (Intelligent Updates, but not LiveUpdates), Trend Micro and Virusbuster
- Beta definitions updates from F-Secure, McAfee, Panda, Symantec and Trend Micro
- To AV companies: There is no participation fee. Feel free to join if you’re not yet included!
Current problems

- From 2004-01-01 until 2004-09-01 we have downloaded more than 37,000 update files
- However, only about 30,000 are “valid”!
- Three main problems:
  - Update servers are out-of-sync (this means, we continuously download old and new updates)
  - Corrupted updates (damaged signatures or archives; e.g. we download a file during an upload process of the AV company)
  - Non-reachable (possible overloaded) servers
- Therefore, all files are sorted manually before use (which is a time-intensive process)
Some interesting general statistics

- In the next few slides you will see:
  - Number of released updates
    (Note: a high number might not be good and a small number might not be good either!)
  - Updates releases of the last few months, per day of the week and per hour
  - Signature update growth rates since 2004-01-01
- Note: all times are in GMT (24 hour format)!
Regular update releases per day (I) (x = date, y = number of updates)
Regular update releases per day (II)

- Days with the most update releases:
  - 2004-03-03 (89), 2004-04-28 (83),
  - 2004-08-16 (80), 2004-03-18 (76),
  - 2004-07-19 (74)

- Days with the lowest number of update downloads:
  - 2004-02-07 (7), 2004-01-04 (8)
Regular update release per company (x = product, y = # of updates / month)
Regular update releases per month (x = number of updates, y = month)
Update releases per weekday
(x = number of updates, y = weekday)

Regular Definitions
Beta Definitions
Update releases per five minutes
(x = time, y = number of updates)
Update releases without Kaspersky
(x = time, y = number of updates)
Signature file growth rates

- Average signature file growth rate of all AV vendors from 2004-01-01 until 2004-09-01 is **24.3%**
- Some examples:
  - AntiVir: 1,616 KB → 1,868 KB (15.6%)
  - Bitdefender: 2,279 KB → 2,839 KB (24.5%)
  - Kaspersky: 3,424 KB → 4,361 KB (27.4%)
  - McAfee: 3,813 KB → 4,606 KB (20.8%)
  - Norman: 1,062 KB → 1,172 KB (10.3%)
  - Panda: 4,872 KB → 6,634 KB (36.1%)
  - Sophos: 5,655 KB → 6,558 KB (16.0%)
  - Symantec: 9,052 KB → 10,688 KB (18.1%)
  - Trend Micro: 7,540 KB → 9,664 KB (28.2%)
Retrospective test results (I)

- Set of 100 different Win32 ItW malware (but not necessarily outbreaks), with variants of the following families:
  - Agobot, Atak, Bagle, Blueworm, Bobax, Evaman, Korgo, Lovgate, Mydoom, Nachi, Plexus, Sasser, Sdbot, Sober, Zafi
- Out of these, 23 were discovered in May, 23 in June, 25 in July, 16 in August, 13 in September 2004
- Scanners were tested in monthly intervals starting at 2004-05-01 until 2004-09-01 (this means, five test-runs)
- Tests show both the heuristic results (in May) and signature-based virus detection, plus the detection development over time
- Note: many virus authors check their new malware against some scanners first, trying to avoid heuristic detection
Retrospective test results (II)
(x = product, y = detection score)
Retrospective test results (III)

- Norman scored best, detecting 39 out of 100 malware proactively using its Sandbox
  - Furthermore, a short analysis of the malware is provided
  - Negative: Requires a lot more scan time
- McAfee scored well, too (30 %), while Trend Micro detected no malware without updates
Outbreak response time test results (I)

- Our starting point (time 0:00 h) = where the first scanner detected the malware with a special (non-generic) signature update
- Proactive detection = response time of 0:00 h, too
- Alternative methods possible, but not used:
  - Starting point = the time where the first sample was seen somewhere in the world or when the outbreak started (but it’s hard to find out the exact times…)
  - Proactive detection = 0:00 h response time, signature detection = a response time of at least 1:00 h
Example: Mydoom.A

- All AV updates which were released on 2004-01-26:
  - F-Prot 22:30 W32/Mydoom.A@mm
  - Trend Micro 22:35 WORM_MIMAIL.R
  - RAV 23:00 Win32/Novarg.A@mm
  - Norman 23:05 MyDoom.A@mm
  - F-Secure 23:05 W32/Mydoom.A@mm
  - Virusbuster 23:05 I-Worm.Mydoom.A
  - AVG 23:15 I-Worm/Mydoom
  - Avast 23:15 Win32:Mydoom [Unp]
  - Kaspersky 23:30 I-Worm.Novarg
  - AntiVir 23:30 Worm/MyDoom.A2
Example: Mydoom.A (continued)

- All AV updates which were released on 2004-01-27:
  - Symantec 00:05 W32.Novarg.A@mm
  - eTrust (CA) 00:20 Win32/Shimg.Worm
  - Command 00:20 W32/Mydoom.A@mm
  - Sophos 00:40 W32/MyDoom-A
  - eTrust (VET) 01:30 Win32.Mydoom.A
  - Esafe 01:50 Win32.Mydoom.a
  - Dr. Web 02:40 Win32.HLLM.Foo.32768
  - McAfee 04:00 W32/Mydoom@MM
  - Quickheal 04:00 W32.Novarg
  - Bitdefender 04:00 Win32.Novarg.A@mm
  - Panda 04:10 W32/Mydoom.A.worm
  - Ikarus 08:35 I-Worm.Mydoom
Example: Mydoom.A (stopped e-mails)
Data source: © 2004 MessageLabs
Outbreak response time test results (II)

- We measured the response times with **publicly available updates** of 45 outbreaks (2004 only)
- Sorry, but we have no results of…
  - ClamAV, because a large number of files in our test set are still not detected (for the detected stuff, mainly e-mail worms, the response time was less than six hours)
  - Fortinet, because the measurement interval was too small (most outbreaks were in the first quarter of 2004, but we started to track Fortinet at 2004-04-15)
  - Esafe, because we don‘t have a working scanner anymore
Average response times (I)

- Less than 2 hours: none!
- Less than 4 hours: Bitdefender and Kaspersky
- Less than 6 hours: AntiVir, Dr. Web, F-Secure, Panda and RAV
- Less than 8 hours: Quickheal and Sophos
- Less than 10 hours: AVG, Command, F-Prot, Norman, Trend Micro and VirusBuster
- Less than 12 hours: Avast and eTrust (CA)
Average response times (II)

- Less than 14 hours: Ikarus and McAfee
- Less than 16 hours: eTrust (VET) and Symantec (Intelligent Updates, but not LiveUpdates)
- Overall response time: about 10 hours
- Note: beta definition update of McAfee (DailyDats) and Symantec (Rapid Release Definitions) were usually available within less than 4 hours
- Many larger AV companies have Service Level Agreements (SLAs) for a predefined response time with special (non-publicly available) signature updates
Average response times (III)

- Reaction times are always a trade-off between a fast response and reliability (think about false positives, non-working or PC-crashing updates)
- The shown number includes only the time for the detection of the main malware component, but not for (possible) dropped files (e.g. keyloggers)
- Another interesting test: Did all companies detect the dropped components with the first update (or with a second update which was available a few hours later), too?
The answer is: NO!

Only 7 out of 24 tested AV companies were able to do it: AntiVir, AVG, eTrust (VET), McAfee, Panda, Sophos and Trend Micro detected everything.

Some companies required a few days to weeks for full detection (not mentioning a full repair).

Thus, AntiVir and Panda had the best complete updates ready within 6 hours! (Anyway, AntiVir doesn’t have incremental updates available yet….)
Response times are a key factor of current AV solutions, but too many updates can be as wrong as too few updates.

AV industry is quite busy: About 25% signature update growth rates for the first 8 months only.

Heuristics of some programs are very good, but consider that the detection is still below 40%.

The overall response time of 10 hours is improvable.

Future tests: IDS and IPS systems (Jan 2005)
Any questions?

- Are there any questions?