1 + 1 ≠ 2
in malware scanning

Taeil Goh
Potentials and pitfalls of aggregating multiple antimalware products

Overcoming the pitfalls
✓ Drawing upon our experience (Metascan with 8 ~ 24 AVs)
Metascan/Metascan-Online

- Metascan
- Metascan-Online
Why Multi-scanning
No AV shows 100% detection

On demand test results from AV comparatives from the 2010 August test and 2011 February test.
Threats missed by one may be found by others

Y axis is percentages of missing detections while 3 or more AVs among 9 AVs reports as threats.
Various response times to outbreak

Comparison of response time to outbreak with selected AVs, AV-Test.org.
Y-axis represents delay time in hours.
Suspicious
Detection ratio

- Heuristic scan increases detection rates
  but also increases false positive
  No easy way to determine false positive

- No more black and white
  Detection ratio (decision making factor)
  www.metascan-online.com
  www.virustotal.com
  virusscan.jotti.org
Fallback/Redundancy

- Software failure in race condition

- Disable temporary proactively
  (e.g., catastrophic update)

- During the regular maintenance of AV such as upgrade
Overcoming Challenge

1+1!=2
Challenge
Downside of more AVs

- Scanning time
  \[ 1+1 > 2 \]

- Potential failure due to an exploit
  \[ 1+1 = 2 \]

- False positive
  \[ 1+1 = 2 \]

Combining downside of AVs:
\[ 1+1 \geq 2 \]
Performance optimization

Strategies to discuss

- Avoid redundant pre-scanning tasks
  Decompressing data

- Reduce scanning needs
  Filtering based on file type

- Avoid redundant scanning
  Caching scan results
Performance optimization
Utilizing File Type Analysis

- Detecting archive file
- Filtering based on file types
  Accept/Reject specific file types
- Multiple algorithms for file type analysis
  E.g., Linux File, TrlD
Performance optimization
Pre-processing Archive File

- Remember
  
  Extracting archive files is very expensive

- Improve detection rate of AV

- Consider
  
  Multiple archive libraries
  Handling bad archive files such as archive bomb recursion level, file size, and file ratio
Caching

carefully

- Remove redundant scanning
  same data is usually seen over and over

- To consider
  Rescan on demand to override cache
  Reset on update of definition database
Summary of performance optimization

For Example

- Scan a file
  - Detect file type
  - Is the file archive file?
    - NO ➔ Reject
    - YES ➔ Extract the file
  - Is file type expected?
    - NO ➔ Reject
    - YES ➔ Continue for scanning
Overcoming Challenge

Downside of more AVs

- Scanning time
  
  \[1 + 1 < 1\]

- Potential failure due to an exploit
  
  \[1 + 1 = 2\]

- False positive
  
  \[1 + 1 = 2\]
Software vulnerabilities

For 9 Advanced+ AVs

Number of secunia advisories on the selected AVs.
Robust integration
Resiliency from minimizing impact

- What we can do?
  - Minimize the impact of AV/components failure

- Multi-process (with Inter-Process Communication)
  - E.g., web browser technology

- Handle DoS vulnerability
  - Timeout for scanning
  - RAMDISK
Robust integration

Watchdog

Diagram:

- Starts monitoring
  - Is process running
    - YES: Is process working
      - YES: Wait
      - NO: Recover the process
    - NO: Recover the process
  - NO: Recover the process

- Alert/Reporting
Comparison of scanning speeds between single process-based solution (marked as Single) and multi-process-based solutions (marked as Multi) for executables (marked as E) and 3788 files without differentiating the file types.
Overcoming Challenge
Downside of more AVs

- Scanning time
  \[1 + 1 < 1\]
- Potential failure due to an exploit
  \[1 + 1 < 1\]
- False positive
  \[1 + 1 = 2\]
Detecting False Positive
Not simple but possible

- No logical OR operation of the scan results
  Utilizing detection ratio
  (e.g., label data as “suspicious” if lower than 25%)

- Integration with comprehensive analysis tools such as sandbox solution.

- Further Manual inspection

- More AVs means
  higher confidence level based on detection ratio
Overcoming Challenge
Downside of more AVs

- Scanning time
  \[1 + 1 < 1\]
- Potential failure due to an exploit
  \[1 + 1 < 1\]
- False positive
  \[1 + 1 < 1\]

Combining downside of AVs:
\[1 + 1 < 1\]
Acknowledgement

Thanks To

AV-Comparatives
AV-TEST
Secunia
OPSWAT

for all the testing results and support of this research
For any question or feedback, please email

taeil@opswat.com