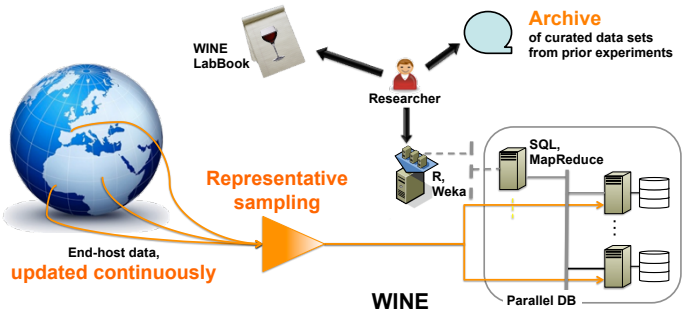
WINE: Data-Intensive Experiments in Security

PI: Tudor Dumitraș, Symantec Research Labs
tudor_dumitras@symantec.com
http://www.symantec.com/about/profile/universityresearch/sharing.jsp

Experimenting with Big Data Ideas

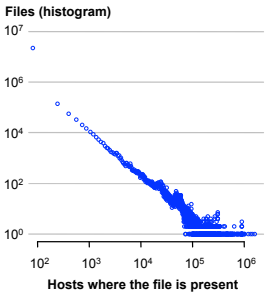
- Big Data is hard to analyze and move around
  - 1 MB on single host or LAN: 0.1–3 ms
  - 1 MB across datacenter: 10 ms
  - 1 MB across Internet: 9,000 ms
- The quality of information is uncertain
  - Field data collected on millions of hosts worldwide
  - Big Data experiments are hard to reproduce
- The data must be representative
  - Security arms race => need updated, Internet-scale data on cyber threats

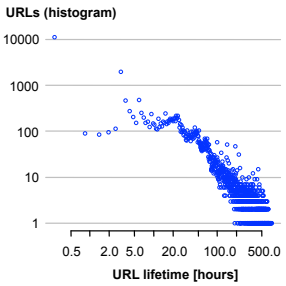


Creating Internet-Scale Models Using WINE

- WINE data set example:
  what executable files do people download?

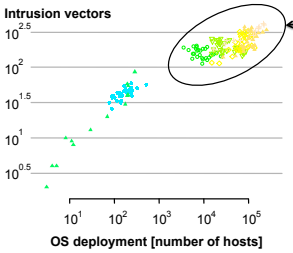
Binary Reputation Submissions
- Machine ID
- Timestamp (client-side & server-side)
- Hash (MD5 &SHA2)
- Download URL



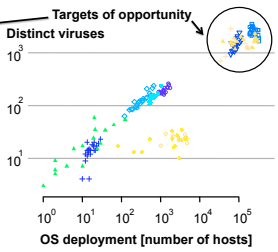


Analyzing Field Data Using WINE [LEET 2012]

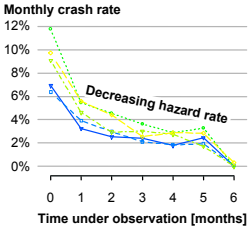
Intrusion-Detection Telemetry



Anti-Virus Telemetry

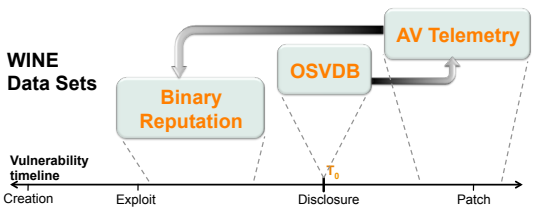


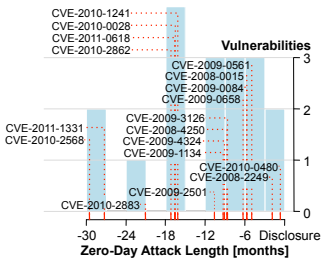
System-Stability Telemetry

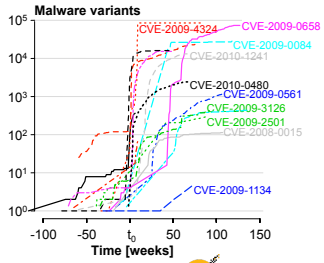


Windows 2000 SP4
Windows 2000 SP3
XP SP2
XP SP3
XP SP1
XP
XP SP2 64-bit
Vista
Vista SP1 64-bit
Vista SP1
Vista SP2 64-bit
Vista SP2
Windows 7 64-bit
Windows 7
Windows 7 SP1 64-bit
Windows 7 SP1

Measuring the Length of Zero-Day Attacks Using WINE [CCS 2012]







National Science Foundation
WHERE DISCOVERIES BEGIN

Symantec.

Interested in meeting the PIs? Attach post-it note below!