Server-side Polymorphism: A Study of a Few Links

Robert Sandilands, Director: Antivirus, Commtouch

Overview

- What is server-side polymorphism?
- How was the data collected?
  - linksmonitor
- Discussing the links
- Conclusions

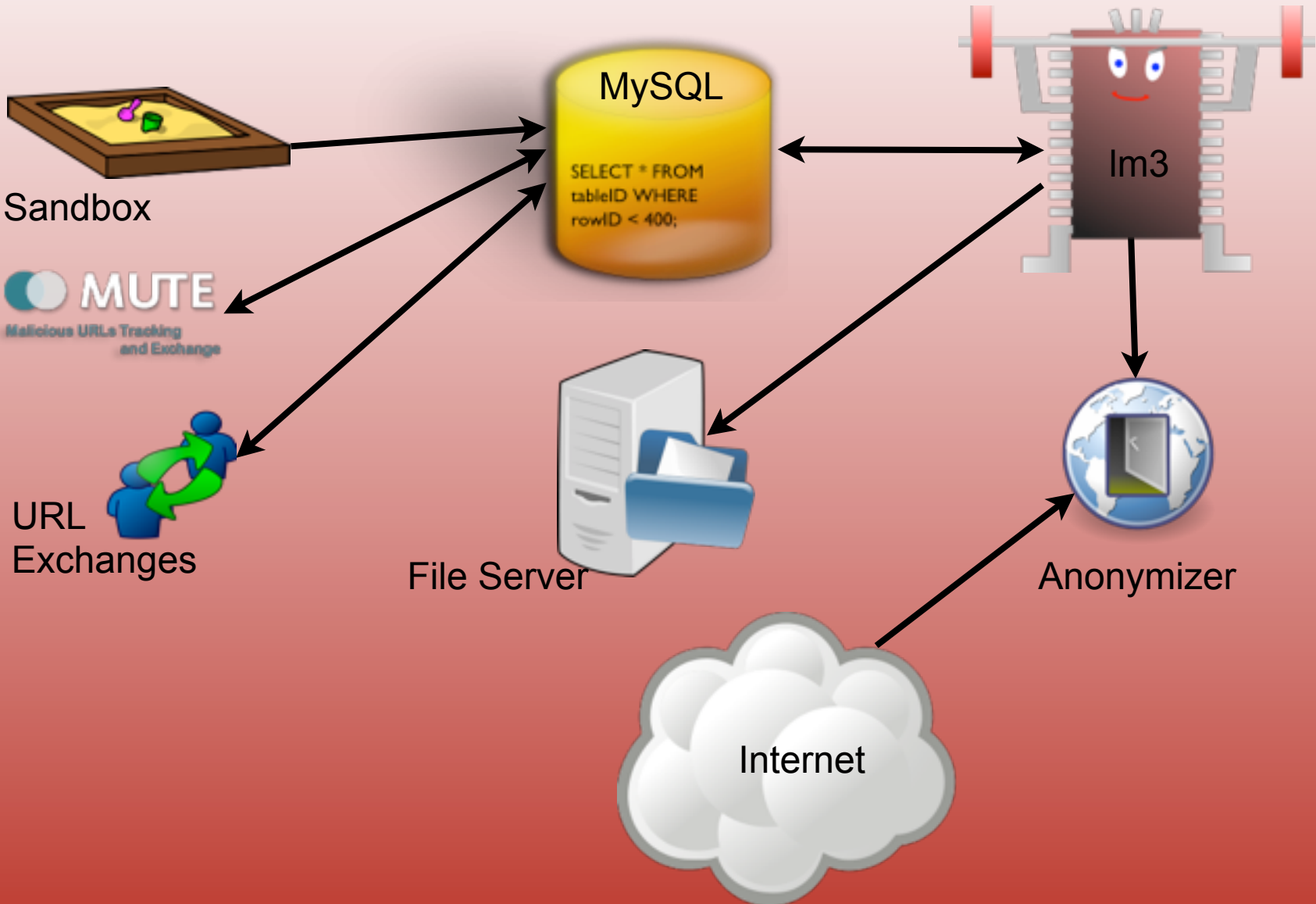
What is server-side polymorphism?

- Polymorphism is mutating code where the functionality stays the same
- Not only code, but also the structure can change
- Downloaded samples (http/https/ftp)
- Changes on a regular basis

How was the data collected?

- Using our linksmonitor system
- Virtual Machine
- MySQL database
- File Server
- C++ Application
  - Multithreaded

Architecture of linksmonitor 3



History of linksmonitor

- Currently third generation
- First generation was a perl script written in 2005
  - Mitglieder
  - Mytob
  - Sober
- We noticed the contents of the download locations changing just before the new malware was spammed out.
- Better response times

History of Linksmonitor 2

- The perl script soon had issues scaling
  - Used local .csv like data file
  - Single threaded
- Rewritten in C in 2006
  - Multithreaded
  - Still using the same .csv files
- Got DDoS'd a few times
- Introduced Anonymizing service
- Started exchanging links

History of Linksmonitor 3

- The .csv file became too unwieldy to handle
- Rewrote it in 2009 in C++
  - uses a MySQL database
- Still multithreaded
- Focusing on server side polymorphism
- Allows us to download active links more often than less active links
- Retires dead links

Limitations of linksmonitor 3

- We get a new IP every day, not every download
- Passive download, no active components
- Focuses on PE executables
- Does not monitor for exploits, scripts, PDF, SWF
- Does not generate usable data for URL Filtering or IP reputation services
- Recently upgraded to 64-bits, but should probably be upgraded to run on a HPC cluster
- Bandwidth limited

Limitations of linksmonitor 3 continued

- False positives
  - regnow.com
  - hotdownloads.com (Digital River)
  - Gambling software

Linksmonitor 3 algorithm

- Every URL has a weight
- The smaller the weight the more often the URL will be investigated
- After every download from the URL the weight is adjusted
  - Lower if something new and interesting was downloaded
  - Higher if nothing new was downloaded
  - Even higher if the URL did not respond

The links

- 5 Links (URLs)
- I will refer to them by my own names which may sometimes match what we detect them by
- For each link I will try to discuss
  - Overview
  - Percentage of samples exchanged
  - Consistency
  - Responses from malware author(s)
  - Nature of changes
  - Naming/grouping

Disclaimers

- Anonymizing all information about competitors
  - I don't want to place people on the spot
  - Glass houses
- Some of this is best guesses and assumptions
- Sometimes the data was just inconclusive
- I assumed heuristic detection if a large number of samples were detected with a similar name
  - Except if prior experience tells me that the name is associated with large groups of unrelated samples
- Some malware names make it obvious it is mass-added

Basic process

- For about 1 month monitored “interesting” links
- Samples collected
- Signatures generated
- Some analysis of the samples were done
- Around 1,500 samples
- Samples submitted to 2 different sandboxes to verify consistent behavior
- Some samples disassembled, but not thoroughly
- Used in-house multiscanner

Link: VB.CK : Overview

- Detected as W32/VB.CK2.gen!Eldorado
- Downloaded 30 unique samples
  - Currently detects more than 1,100 samples
- Sizes varied between 4,204 and 176,781 bytes
- Downloaded a mix of ZIP and PE-EXE files
- VB6 based malware
- Initially focused on 1 URL
- Found 2 more URLs serving the same malware
- Around 36% of samples were also received through sample exchanges

Link: VB.CK : URL

- Primary link was up 2011-12-10 to 2012-03-20
- All URL's seem to be Brazilian
- Probably Brazilian banker malware
- Received the link through a URL exchange with a competitor so the link may have been up earlier
- Malware author had a naming standard for his malware
  - Rainha.Do.Anal.Scr
  - Sexo.Anal.Delicia.Scr
  - ...

Link: VB.CK : Functionality

- The basic functionality of the malware was consistent between samples.
- It downloaded an executable
- Saved it in a new folder created in C:\
  - PerfilLogis
  - Skype
  - debug
- Some possible C&C type traffic was seen
- First versions used direct download, later versions used URL shortening services

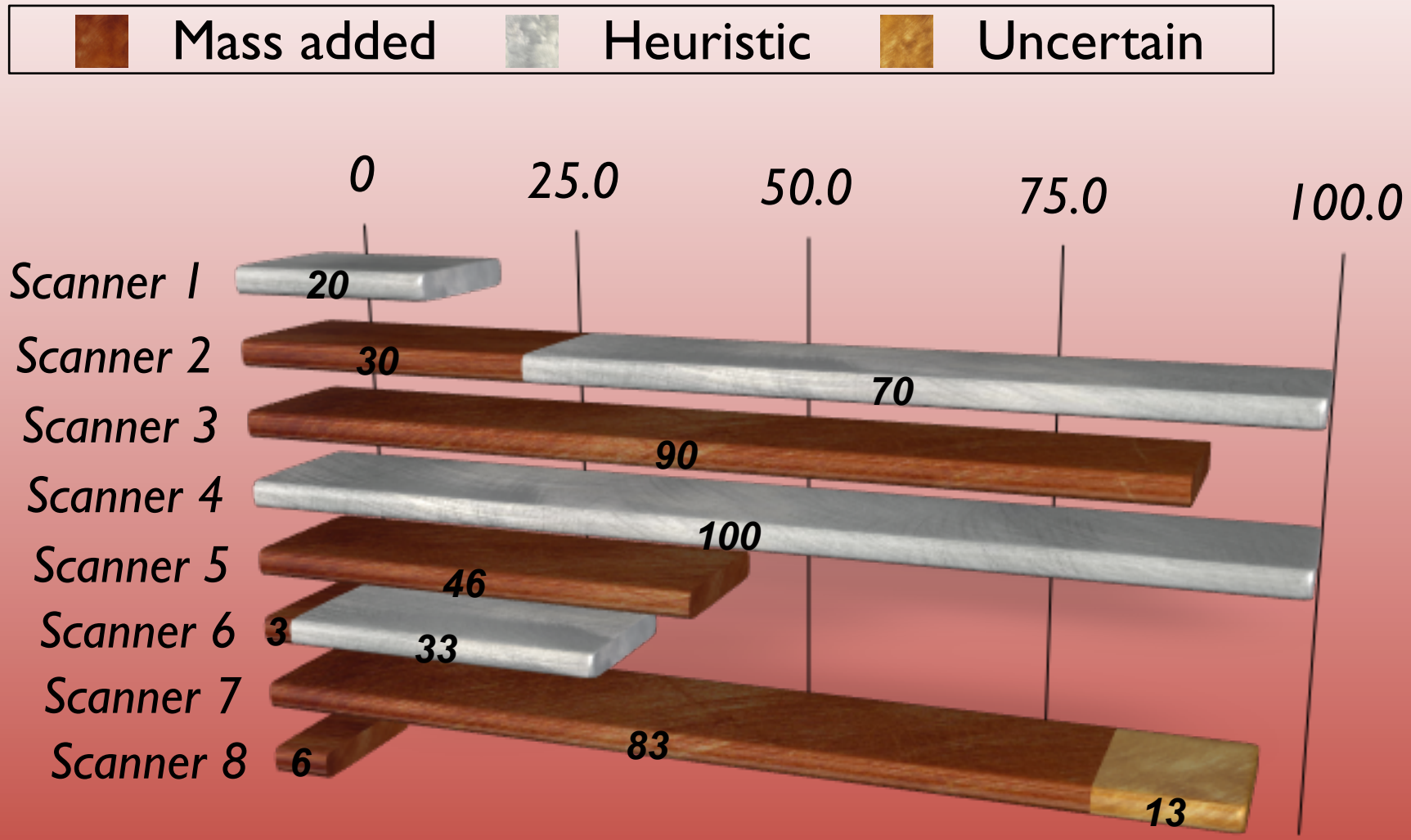
Link: VB.CK : Response

- The malware author was oblivious
- For 1 day (2012-03-15) we detected the malware with a signature that was very old
- Changes in the malware seems to be related more to very active development than real polymorphism
- New versions were released every few days
- Individual samples could be up for anything from a few hours to several days

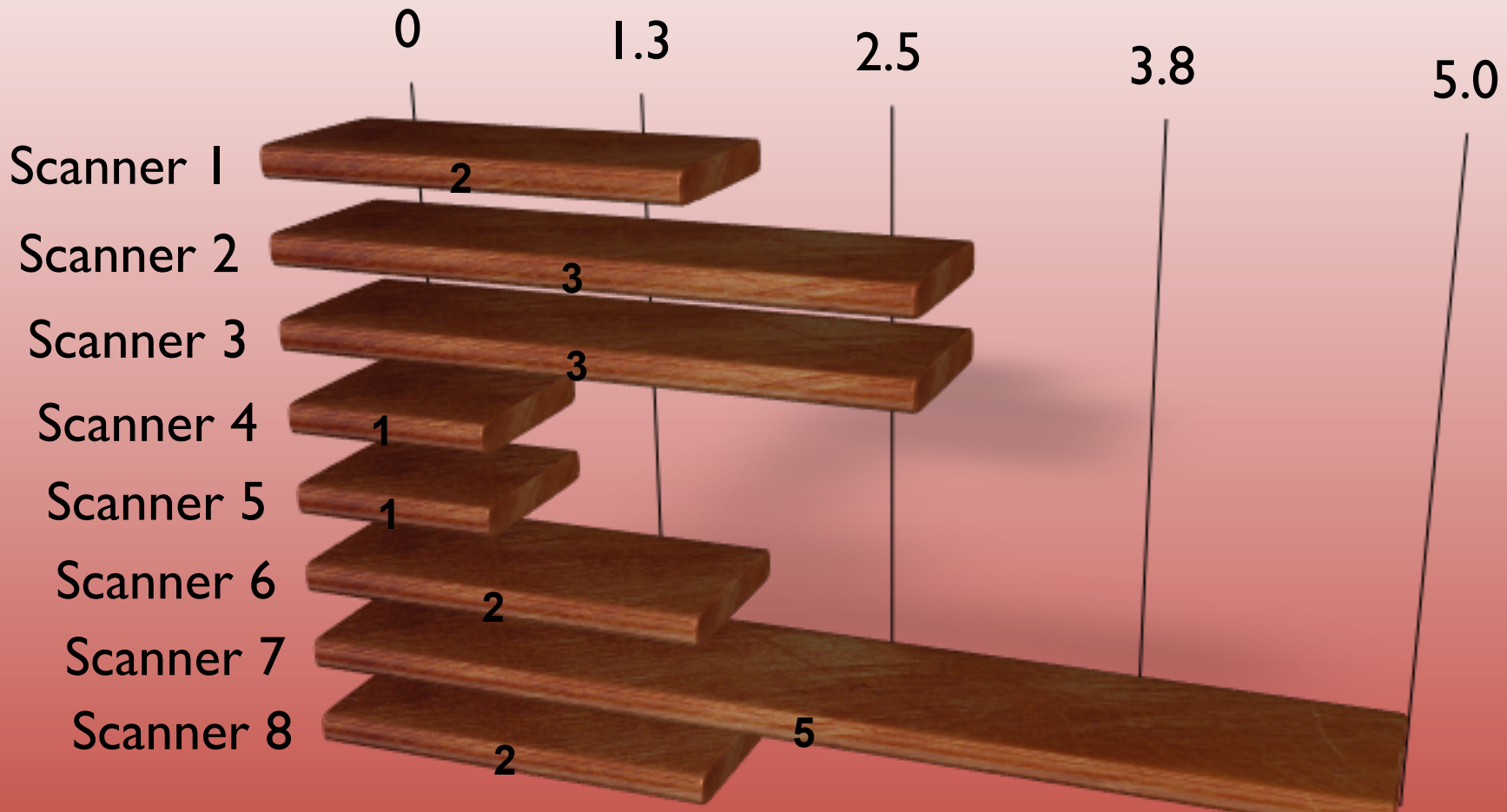
Link: VB.CK : Changes

- The biggest change between versions was the download location or URL shortening service used
- In some versions the PE EXE were also inside a ZIP file and others not
- Could not see any signs of polymorphism
- Changes seems to be related to active development

Link: VB.CK : Detection



Link: VB.CK Unique Family Names



Link: Rorpian.B : Overview

- Probably the most interesting of the links
- Custom packer
- Initially focused on 1 URL
- Found 16 other URLs serving the same malware
- Most of the additional URLs from MUTE
- Originally detected as W32/Rorpian.B.gen!Eldorado
  - Released B2, B3, B4, B5, B6 and B7 to deal with new variants
- Started with 149 samples
- We currently detect more than 7,000 samples with this signature

Link: Rorpian.B Overview 2

- File sizes between 173,568 and 315,065
- Around 51% of samples were also received through sample exchanges
- Anti emulation tricks
  - Testing valid and invalid parameters to function calls
- Probably written using Visual C++
- Probably FakeAV related

Link: Rorpian.B URL

- Primary link first seen 2011-11-10
- Still active at the time of writing
- Most of the other links were alive between 5 days and 2 weeks
- Primary URL registered to somebody in Woodstock, OH
- The web-site does not seem to have any other content

Link: Rorpian.B Functionality

- Behavior of the samples were consistent
- DNS hardcoded to 8.8.8.8
- Stops Windows Security Center Service
- Registers a service
- Windows PE Executable, not a DLL
- Exports functions
- Uses ADS
- Injects winlogin.exe

Link: Rorpian.B Response

- Custom packer and a limited number of samples
- We constantly had to tweak the heuristics to detect new samples
- Unfortunately it was hard to prove whether the changes were just part of the natural variation in the malware or a response to our signatures
- It seems some of the changes may be related to being detected

Link: Rorpian.B Changes

- Random imports
- Random exports
- Random file sizes
- Random section sizes
- Random number of sections
- First and last section always had the same name
  - .text and .reloc
- Other sections had random names
- 4 to 12 sections

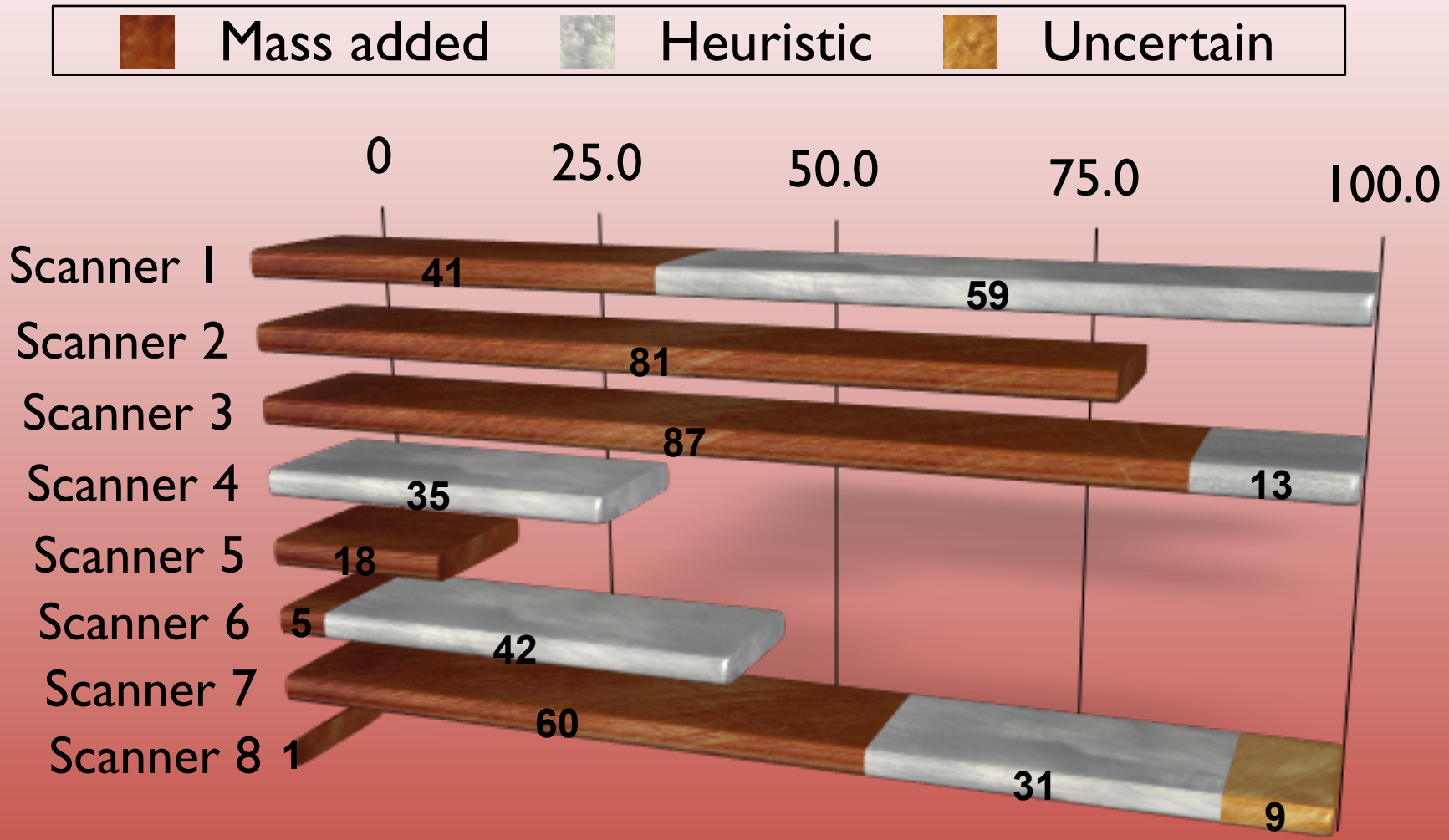
Link: Rorpian.B Changes 2

- Very large number of junk instructions
- Very large number of useless jumps
- The .text section was by far the biggest section
- Custom packer
- Function calls used to detect emulation varies considerably
- Most values in the PE header is modified with each version

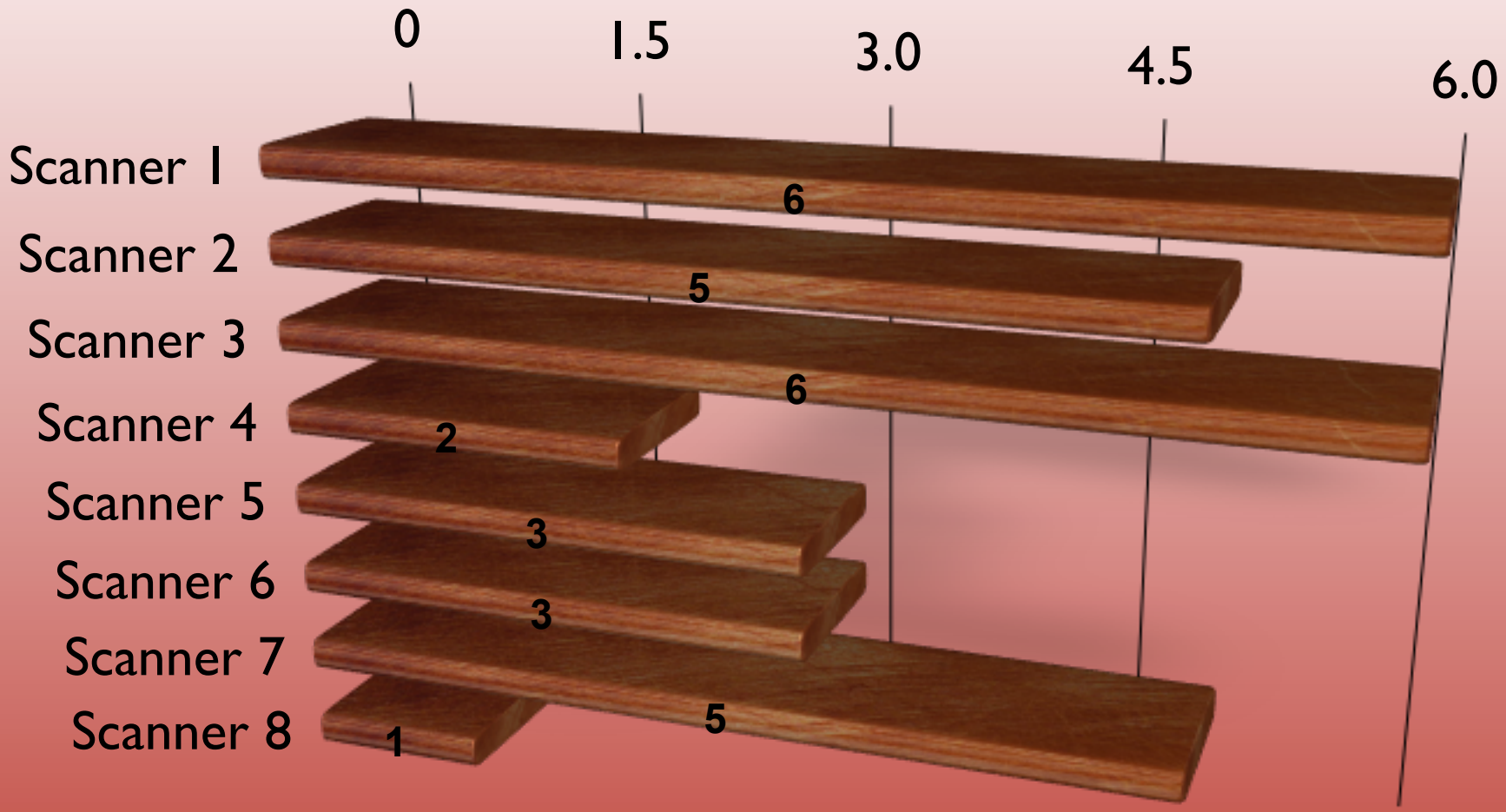
Link: Rorpian.B Changes 3

- Downloaded sample seems to change every hour
  - Assume some automated repacking or rebuilding
- Probably true server side polymorphism
- Could be worse, you could be getting a new sample every time you hit the site

Link: Rorpian.B Detection



Link: Rorpian.B Unique Family Names



Link: Tibs.AC Overview

- Samples were originally detected as W32/Tibs.AC.gen!Eldorado
- Released AC2 to handle changes
- Initially focused on 2 URLs
  - Clustered because of similar php script name
- Found another 31 URLs sharing the same malware
  - Most of the URLs were from a very unreliable source
- Initially worked with 66 samples
  - Signature detects around 1,700 samples
- Around 32% of samples were also received through sample exchanges

Link: Tibs.AC URL

- Initially seen 2011-12-31
- Last seen 2012-03-21
- Some URLs served only a single sample for several weeks
- Some URLs served a new sample every hour
- One of the URLs served 4 completely unrelated samples in the middle of serving Tibs.AC.
  - These samples probably FakeAV
- Some used a php script
- Some directly downloaded the executable

Link: Tibs.AC Functionality

- Except for the 4 unrelated samples all the samples behaved consistently
- Tried to download additional executables from other sources

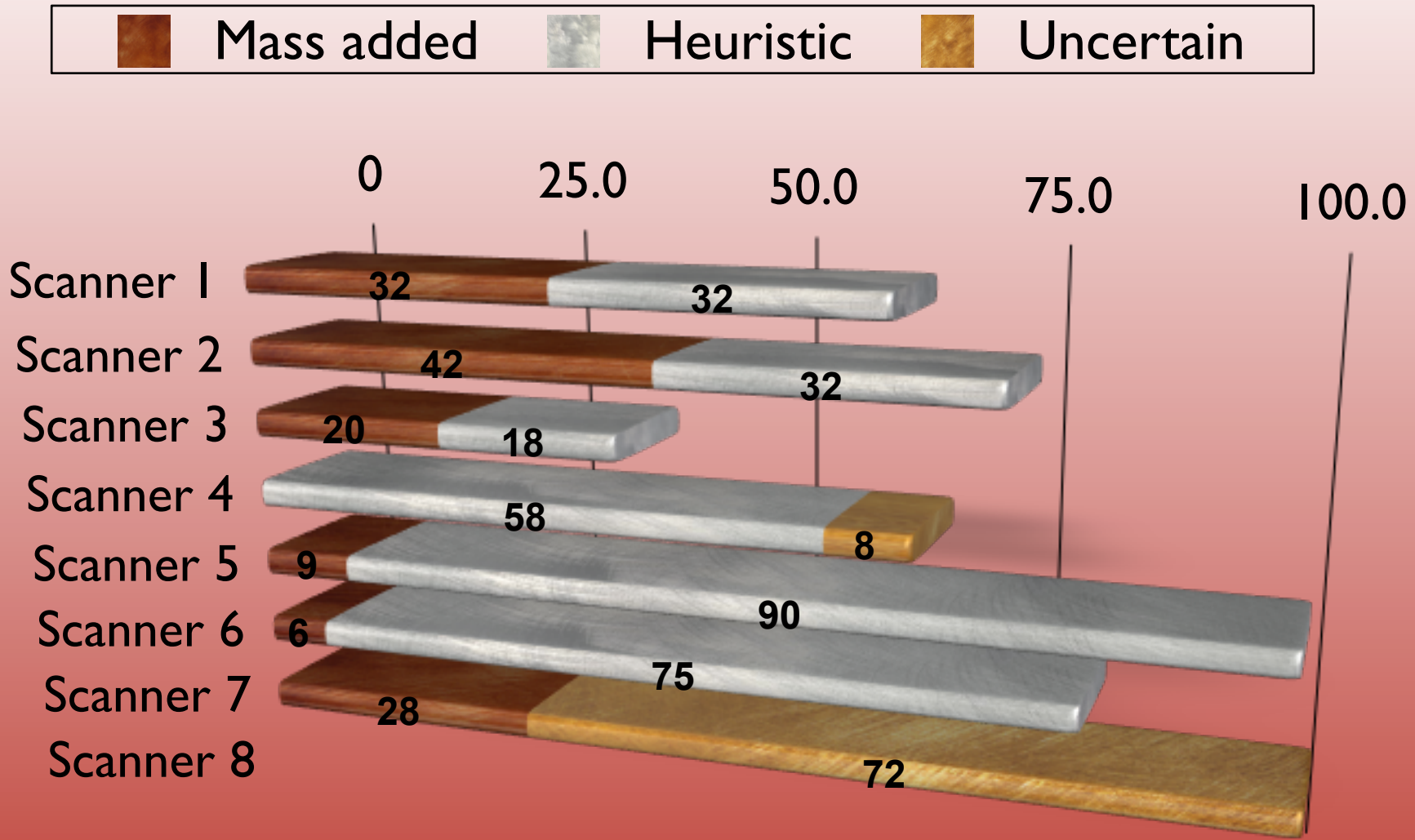
Link: Tibs.AC Response

- There did not seem to be any response or awareness on whether we detect the malware or not
- None of the links were up for long

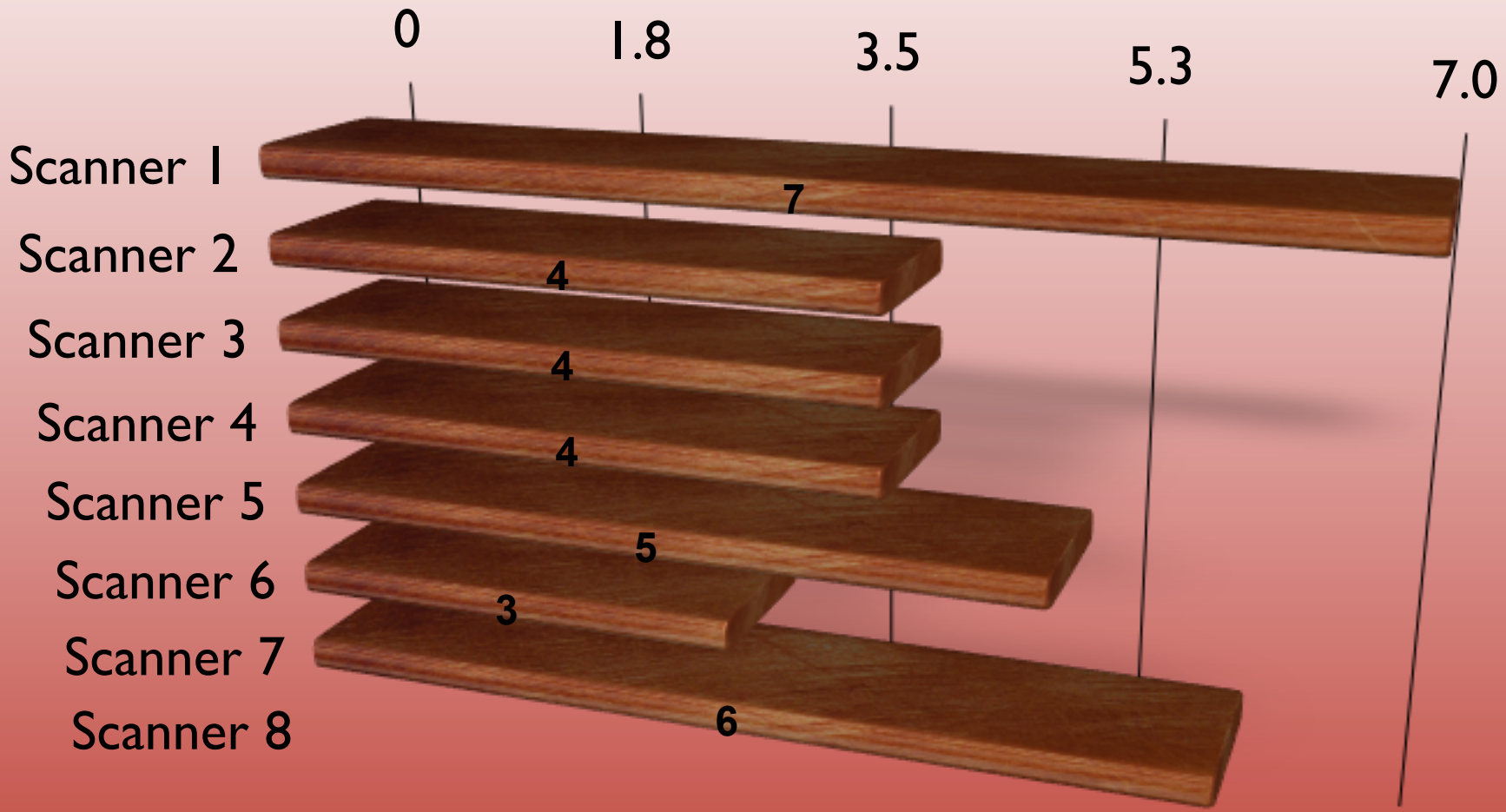
Link: Tibs.AC Changes

- Random imports
- Dynamically imports randomly named DLL
- Variable number of imports
- Section names always the name
  - .text, .rdata, .data, .rsrc, .reloc
- Section sizes different
- File size range: 34,816 to 62,464 bytes
- Custom packer

Link: Tibs.AC Detection



Link: Tibs.AC Unique Family Names



Link: Menti Overview

- Initially looked at 3 links.
  - Same site, just different locations
  - Found another 10 links
  - Mostly hxxp://xxx.xxx.xxx.xxx/files/1[79]
- Modified existing signature to add detection for the links
- Detect as W32/Menti.B4.gen!Eldorado
- A signature similarity test highlighted the samples as being close to the signature that got modified

Link: Menti Overview 2

- Received 5% of samples through sample exchanges
- Initially worked with 88 samples
- Currently detect more than 2,600 samples

Link: Menti URL

- First seen 2012-03-18
- Last seen 2012-04-12
- Most links up for less than 1 day
- The longest one was up for one week
- Some links would serve the same sample if polled very quickly
- Most links seem to serve a new sample at least every hour, but it probably changed more often than that

Link: Menti Functionality

- Detected one of the sandboxes so no useful output
- In the other sandbox it just seems to do a few posts
- Drops some temporary files
- Behavior consistent for all the samples
- Seems to use Gnutella P2P network

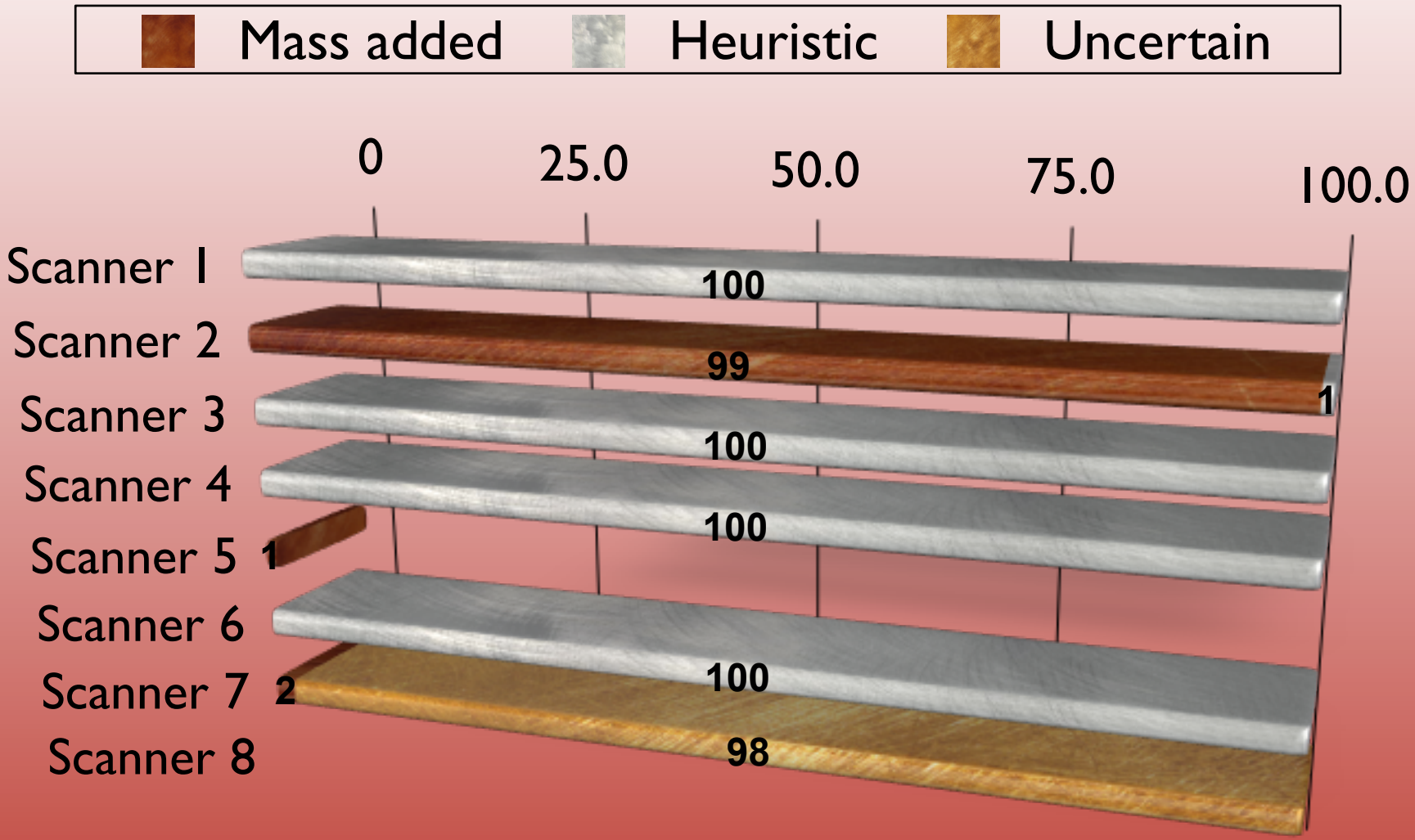
Link: Menti Response

- Did not seem to respond to us adding detection
- None of the download locations were up for very long

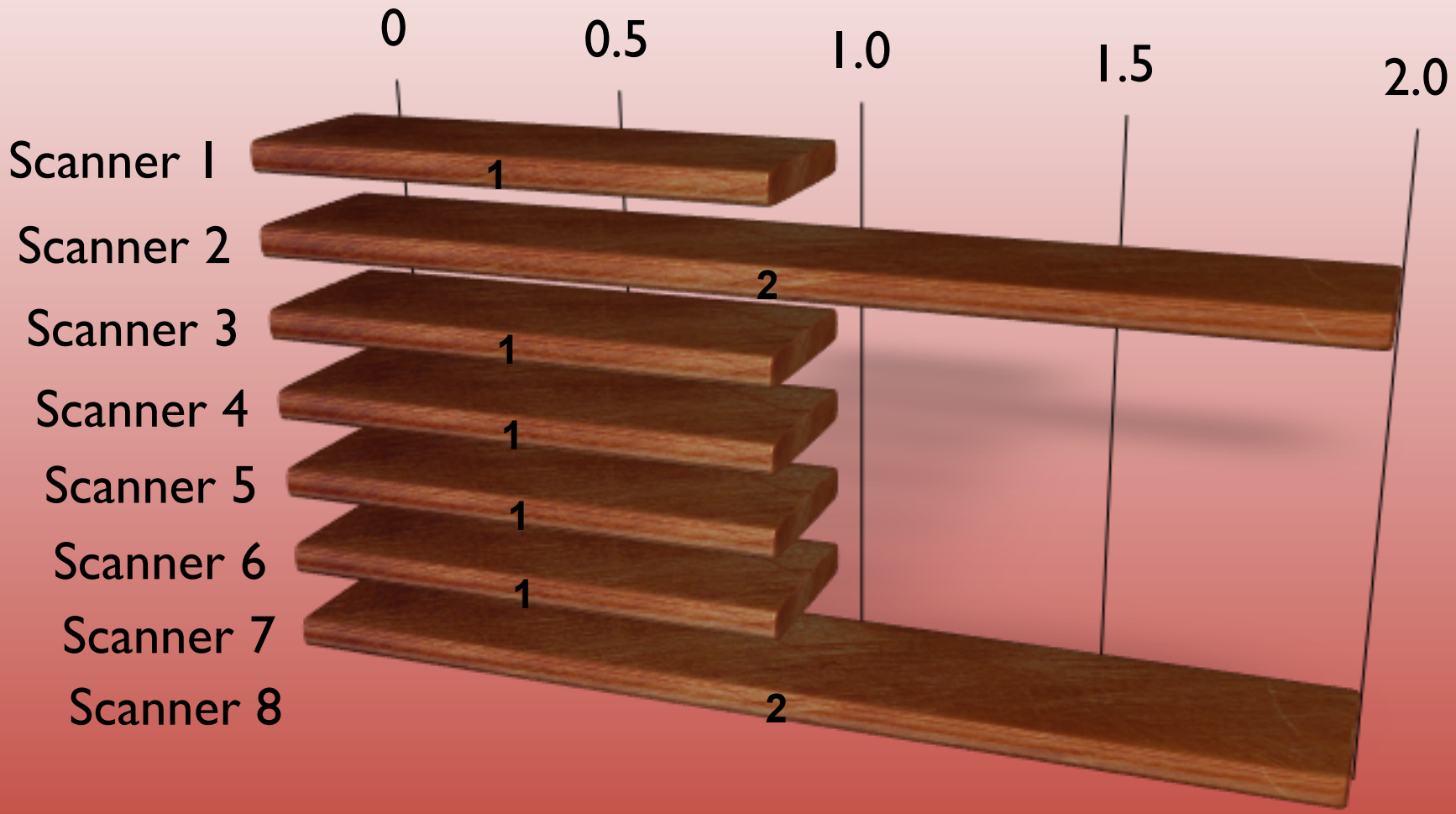
Link: Menti Changes

- File size constant
- 2 sections, same name, same size
  - .text, .rdata
- The only change seems to be some random data inserted in the first section before the actual code starts
- Actual code seems to be the same

Link: Menti Detection



Link: Menti Unique Family Names



Link: Comisproc Overview

- The name is wrong, but this is what I called it in the beginning so it stuck
- Released detection as W32/TrojanClicker.G.gen! Eldorado
- Brazilian malware
- Originally worked with 27 samples
  - Currently detecting around 90 samples
- 37% of the samples received through sample exchanges

Link: Comisproc URL

- First seen: 2012-02-25
- Last seen: 2012-03-19
- No other similar links identified yet
- URL received from public source
- URL points to ZIP file
- ZIP contains UPX packed executable
  - Except for 4 samples
    - BeRo!
- Most samples downloaded multiple times

Link: Comisproc Functionality

- Two different malware families downloaded from 1 URL
- Family 1 was packed with BeRo!
  - Could not get them to run in a sandboxed environment
  - Probably FakeAV
- Family 2 was UPX packed
  - Drops .cpl file
  - .cpl file executed via rundll32.exe
  - .cpl file opens IE to a Brazilian Big Brother site
  - .cpl file BeRo! packed
- File name in zip video[random letters].(exe|cpl)

Link: Comisproc Response

- Detection was only released after the link was taken down
  - Unable to comment on whether there was any response or awareness by the author on whether the malware was detected or not
- Overall detection was pretty low but that could be due to the low risk posed by this malware

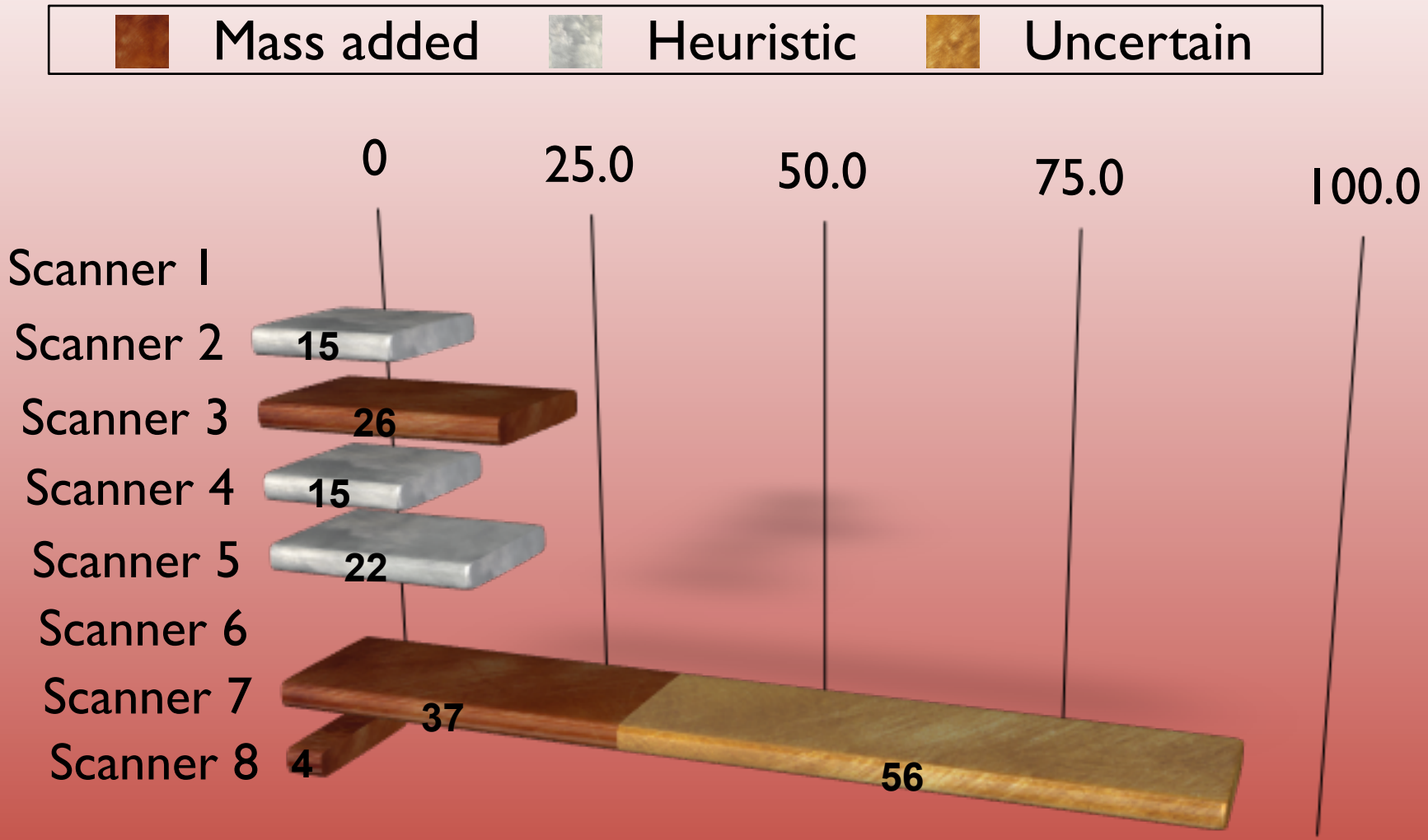
Link: Comisproc Changes

- Code of the dropper seems to be the same
- Borland based compiler
- Variable sized text blocks seems to be the only change between versions
- Some of the text blocks referenced from code
  - Probably an encoded copy of the dropped .cpl file
- ZIP file between 129,477 and 131,316 bytes
- Most EXE files are 136,704 bytes

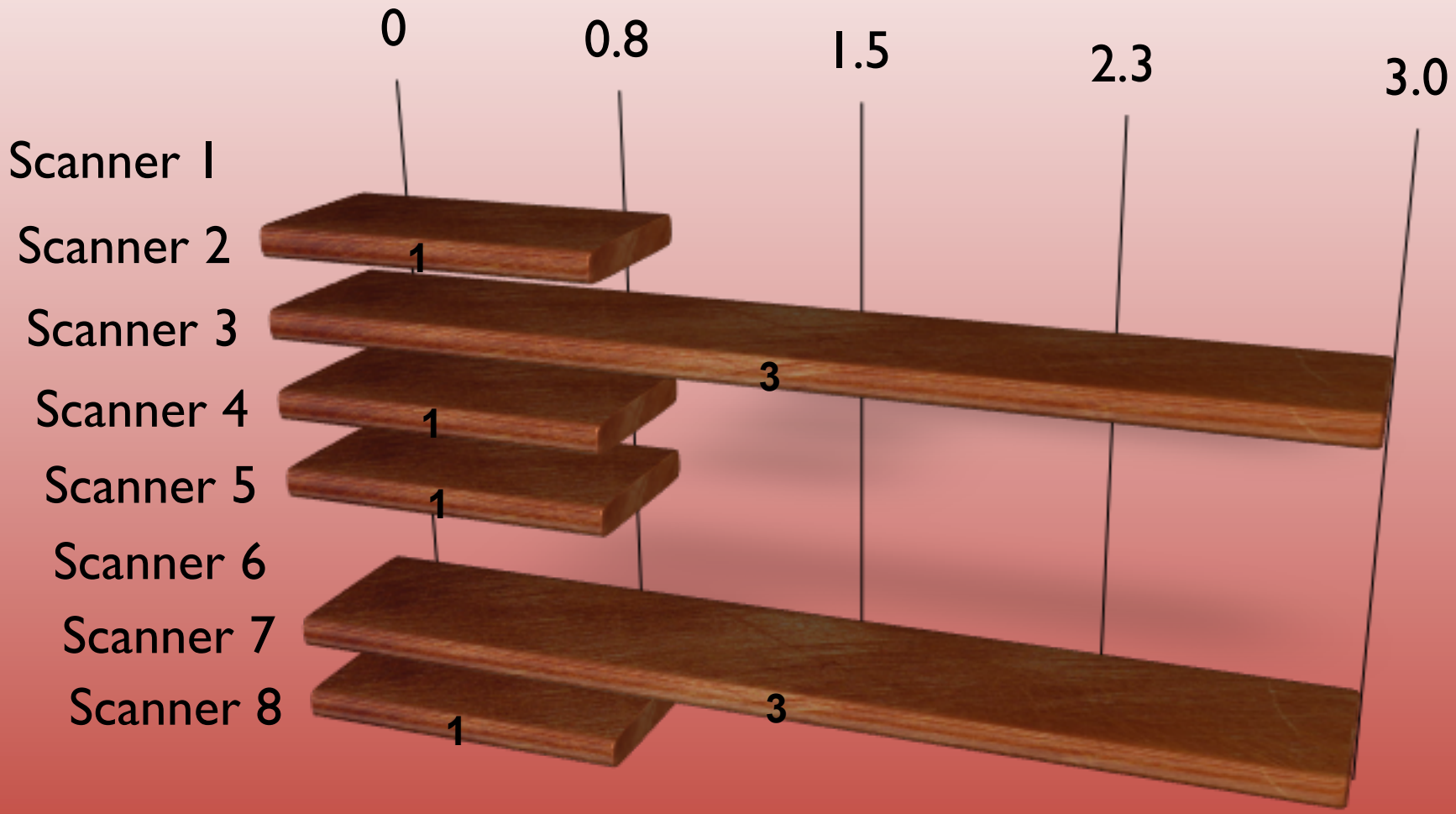
Link: Comisproc Changes 2

- Most changes probably due to changes in code in the dropped .cpl file
  - Slow rate of change
- Changes seems very significant on the highest level due to multiple levels of encoding
  - BeRo!
  - Text
  - UPX
  - ZIP

Link: Comisproc Detection



Link: Comisproc Unique Family Names



Link: Comisproc Detection 2

- My very first impression was that scanners 2 and 4 had detected some samples with some lucky heuristics
- In the end it turned out that two different malware families were served from the same link and one was well detected by those scanners and the rest had bad detection overall
- That is assuming that you agree that opening a specific page on the Brazilian Big Brother site is malicious.

Conclusions

- The links covered a variety of “server-side” polymorphism.
  - Less than expected
  - Very active development
- Sometimes simpler methods seems to be more effective at avoiding detection than complex methods
- Heuristics vs. apparent mass-adding
  - Inconclusive
  - Does mass-adding actually help customers?

Conclusions 2

- Scale out of signatures
  - Only detecting thousands
- Our best heuristics detects more than 90,000 samples per signature

Questions?

Robert Sandilands, Director: Antivirus, Commtouch

Email: roberts@commtouch.com