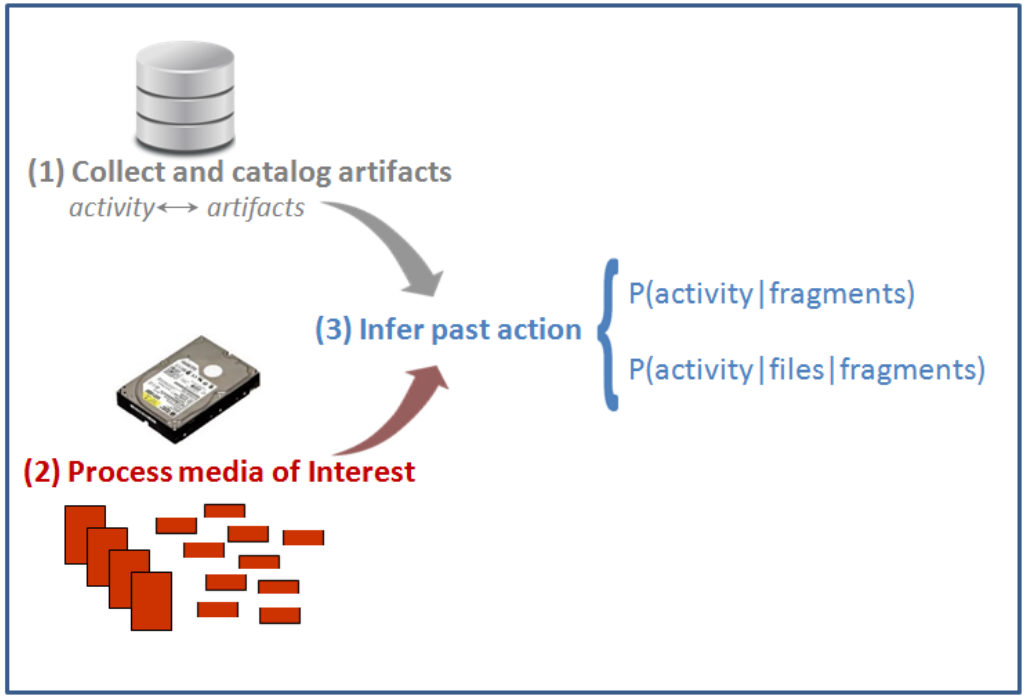
Inferring Past Activity from Partial Digital Artifacts

Jim Jones[†], Tahir Khan[†], Kathryn Laskey[†], Alex Nelson[‡], Mary Laamanen[‡], Doug White[‡]

[†]George Mason University, [‡]National Institute of Standards and Technology



OSDFCon 2015

This presentation results from research supported by the Naval Postgraduate School Assistance Grant/Agreement No. N00244-13-1-0034 awarded by the NAVSUP Fleet Logistics Center San Diego (NAVSUP FLC San Diego). The views expressed in written materials or publications, and/or made by speakers, moderators, and presenters, do not necessarily reflect the official policies of the Naval Postgraduate School or the National Institute of Standards and Technology, nor does mention of trade names, commercial practices, or organizations imply endorsement by the U.S. Government.

A user may uninstall an application to disguise past usage

Malicious Windows user installs wireshark...

...sniffs network, captures account passwords...

...deletes pcap file, uninstalls wireshark...

...uses account info to commit a crime such as impersonate another user or access sensitive information...

...law enforcement confiscates computer...

...can we show wireshark was used in the past?

We reason over partial file artifacts to infer past application usage

Situation:

Uninstalling an application deletes files associated with the application. These deleted files decay over time, i.e., pieces (sectors) of the deleted files are overwritten. Current forensic techniques rely on finding whole and intact deleted files, which may not be available.

Question:

Can we infer past application installation and use when the application has been uninstalled and activity such as reboots and normal usage have continued?

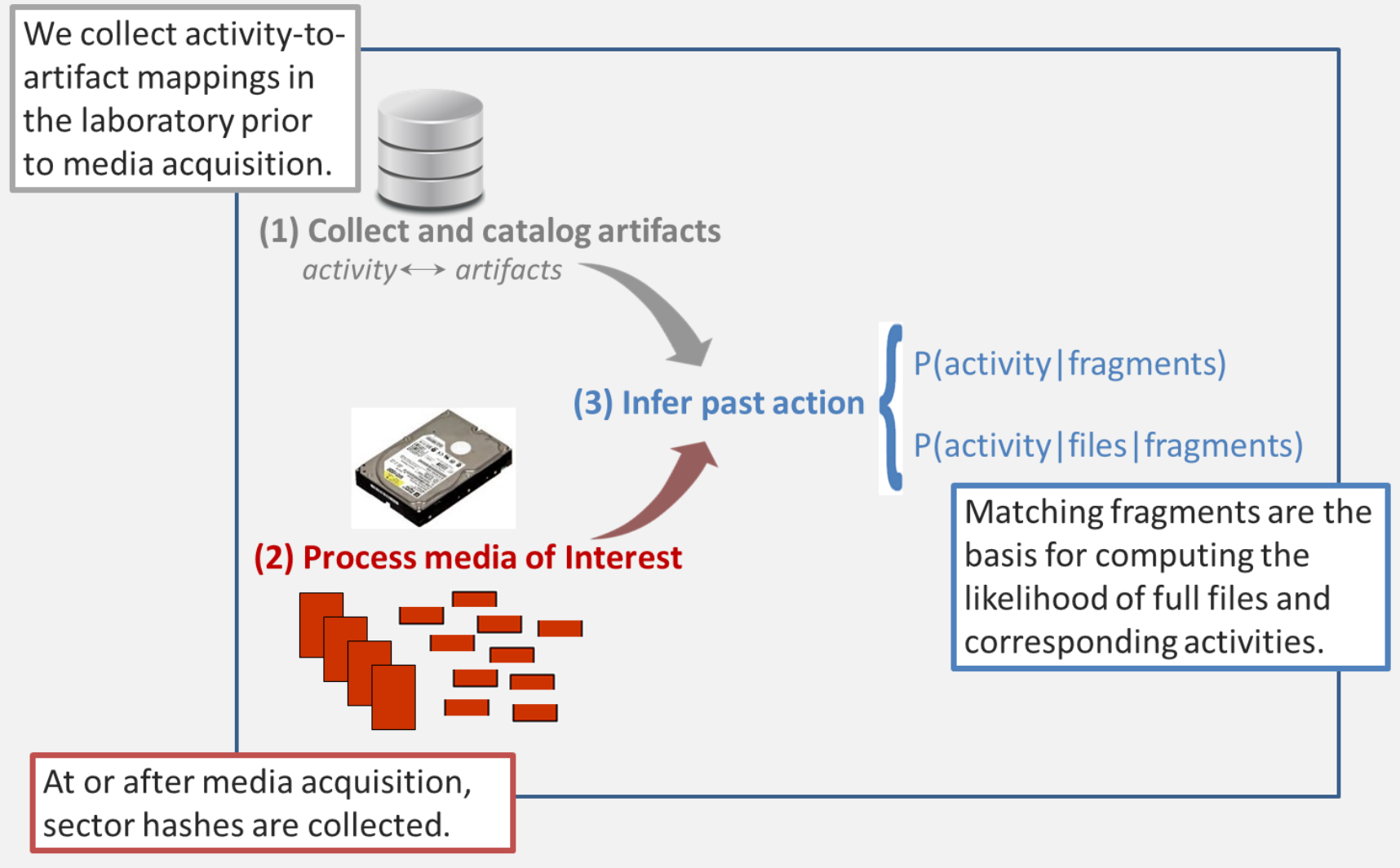
Answer:

Yes, by reasoning over the artifact fragments (file sectors) that remain.

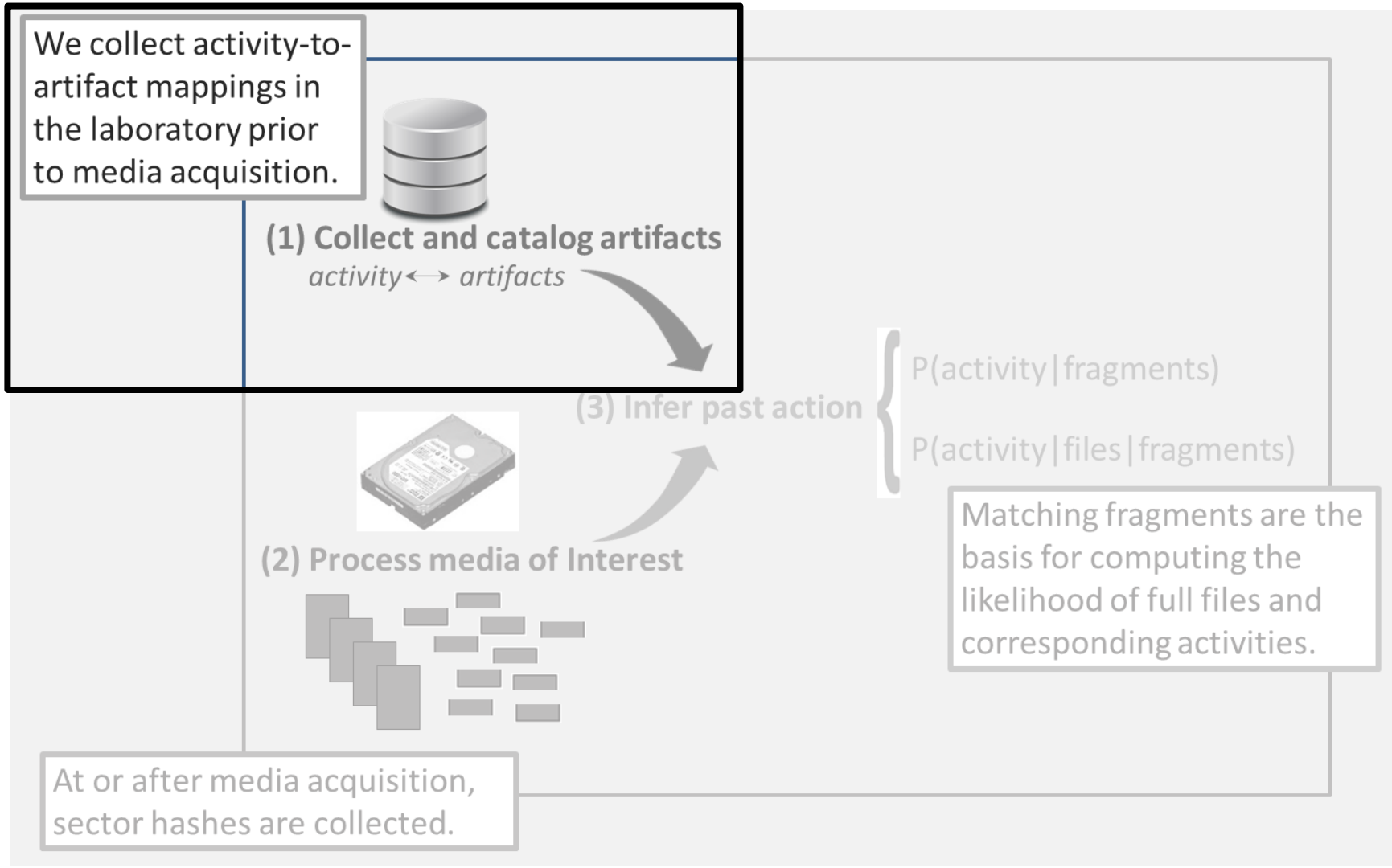
Innovation:

Reasoning over *weighted collections* of artifact fragments.

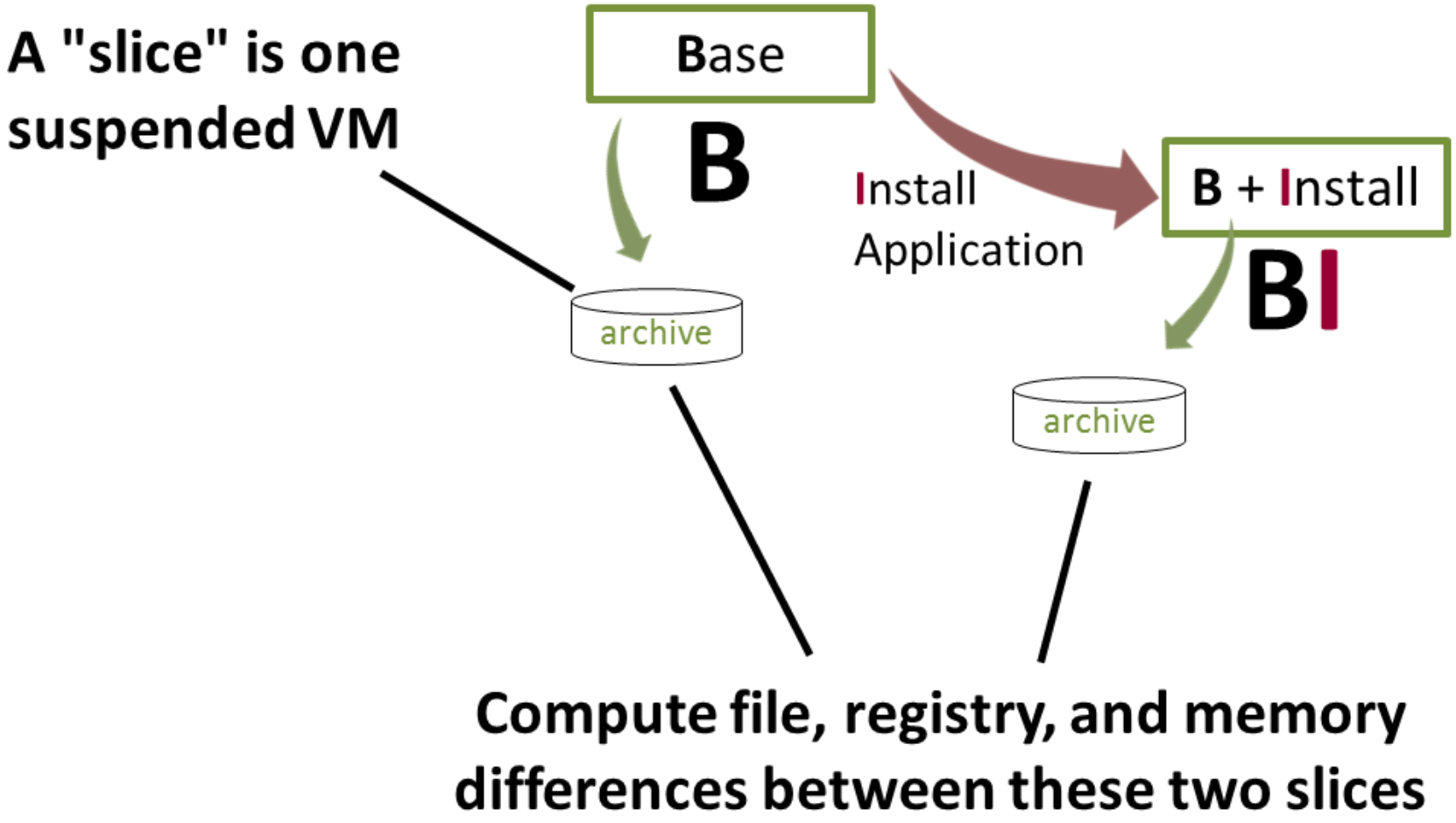
Our approach reasons over media sectors that match a database associating sectors with application activity



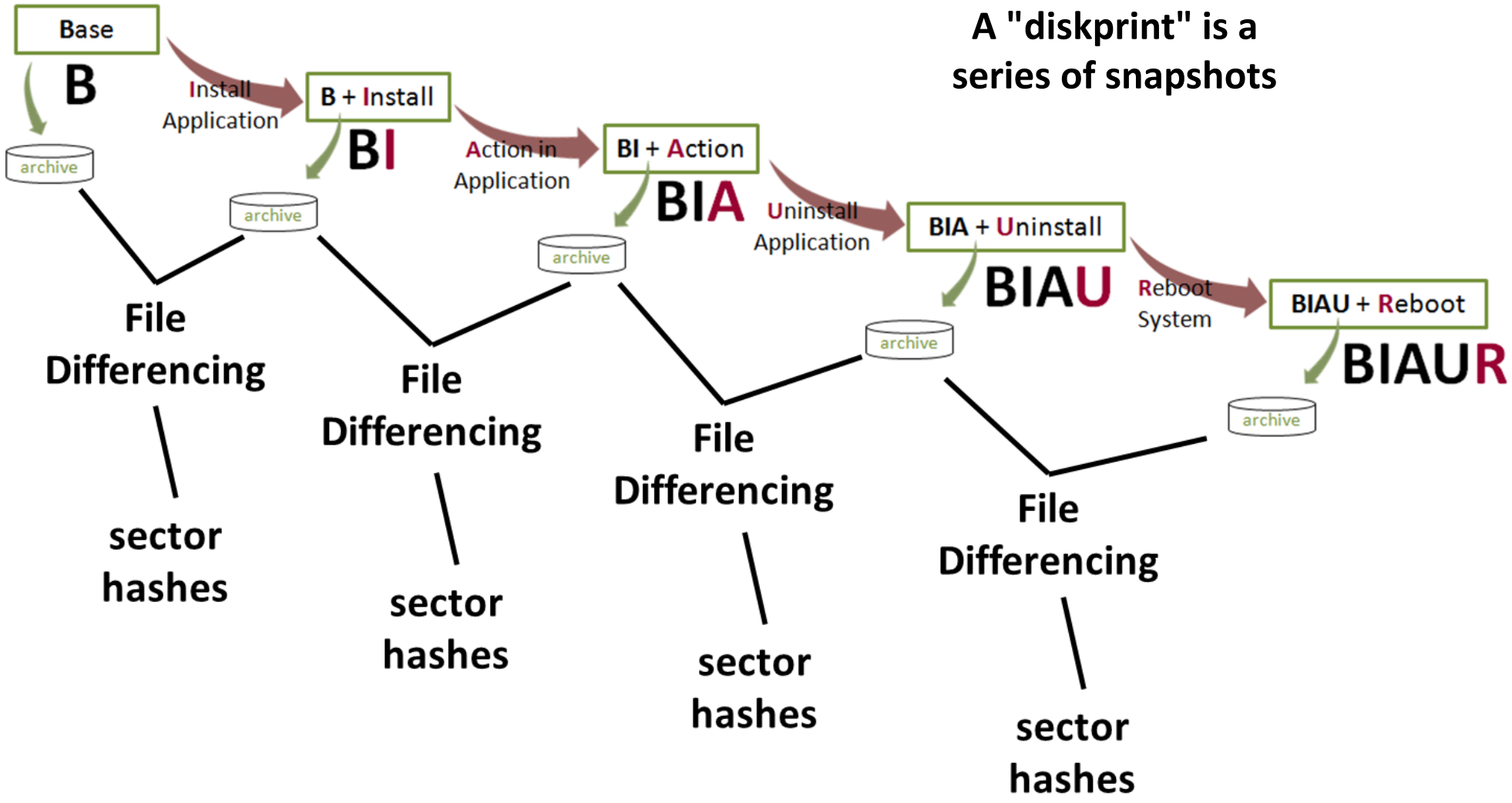
Step 1: Collect and catalog artifacts



We collect activity artifacts using file differencing



We repeat the file differencing process to collect artifacts (files) over a sequence of related activities



Initial diskprinting generated 93M sector hashes from 66k files

16 applications:
- Adv Keylogger
- Chrome
- Eraser
- Firefox
- HxD hex editor
- Invisible Secrets
- MS Office
- Python
- Safari
- Sdelete
- Thunderbird
- TrueCrypt
- UPX
- WinRar
- WinZip
- Wireshark

3 operating systems:
- Windows XP (32 bit)
- Windows 7-32bit
- Windows 7-64bit

4-6 actions:
- Install
- App Activity:
  - Open
  - Use
  - Close
- Uninstall
- Reboot

Data set:
- 29 diskprints
- 186 slices
- 167 difference sets
- ~66k files
- ~93M hashes
  - f < 100

	WinXP	Win7x32	Win7x64
Adv Keylogger	✓		
Chrome	✓	✓	✓
Eraser		✓	
Firefox	✓	✓	✓
HxD hex editor		✓	
Invisible Secrets	✓		
MS Office	✓	✓	✓
Python	✓		
Safari	✓	✓	✓
Sdelete		✓	✓
Thunderbird	✓		
TrueCrypt	✓		
UPX		✓	✓
WinRar		✓	✓
WinZip		✓	✓
Wireshark		✓	✓

We remove file differencing noise and non-probative artifacts

Three categories of artifacts are collected:

A. spurious

B. positively attributed but not probative

C. positively attributed and possibly probative

Select category C by post-processing:

- include by keyword (owning file's path and filename)
- exclude by OS image comparison
- exclude if low entropy
- include if frequency < 100

RESULT: ~8M hashes from ~20k files stored in a hashdb instance

Keyword whitelisting, high frequency sectors, and final hash and file counts

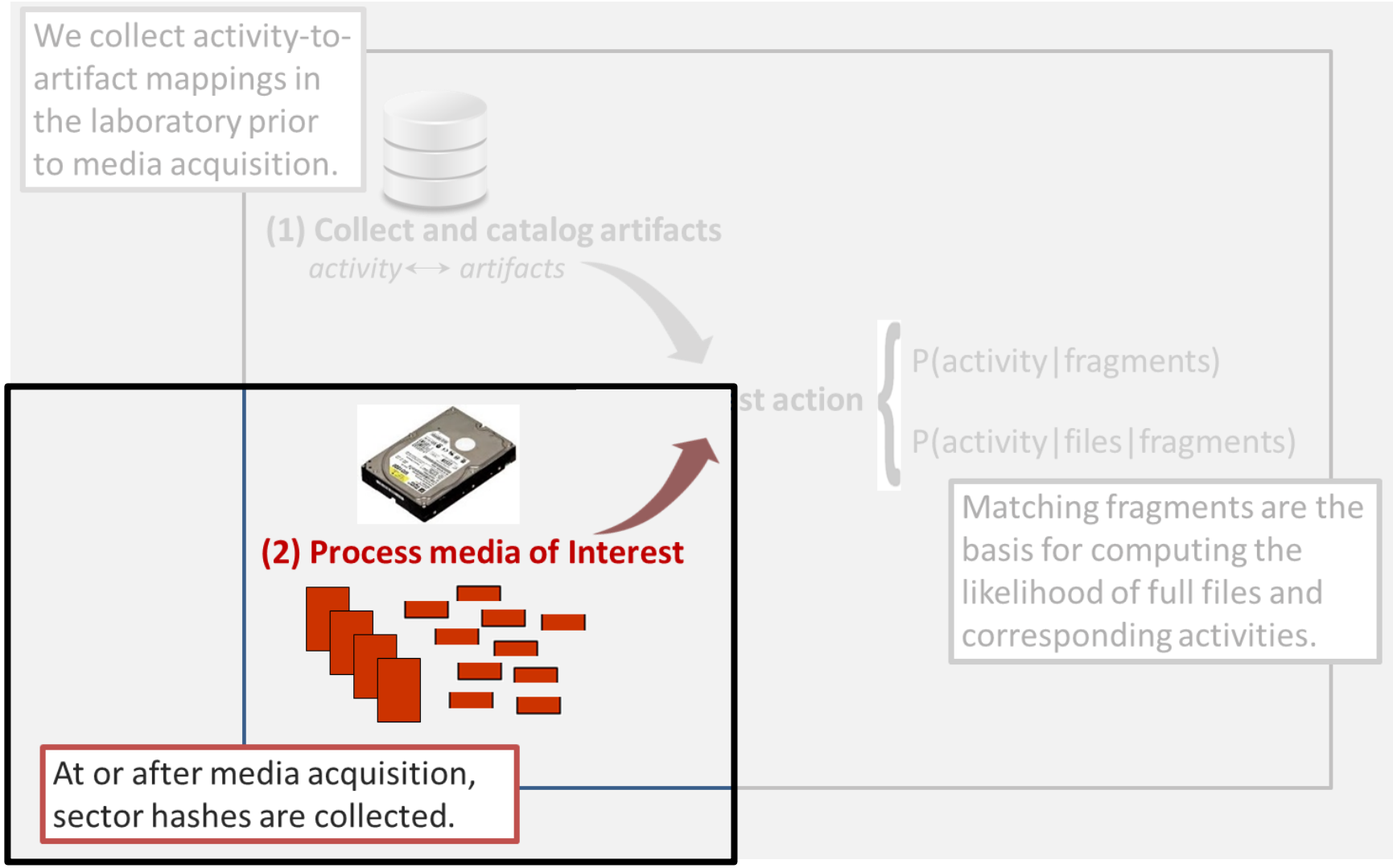
Application	keyword(s)
Adv Keylogger	keylogger
Chrome	chrome,google
Eraser	eraser
Firefox	firefox,Mozilla
HxD hex editor	hxd
Invisible Secrets	"invisible secrets"
MS Office	office,"microsoft shared"
Python	python
Safari	safari
Sdelete	sdelete
Thunderbird	thunderbird
TrueCrypt	truecrypt
UPX	upx
WinRar	winrar
WinZip	winzip
Wireshark	wireshark

High frequency, low entropy sectors (MD5 hash values):

```
bf619eac0cdf3f68d496ea9344137e8b   # repeated 00
393a0fa0f348fb03871ab93726057ddc   # repeated 01
de03fe65a6765caa8c91343acc62cffc   # repeated FF
c5d77850e62433f25d5496bfad94c1b2   # repeated 00 w/
                                   # 06 at offset 510
```

Diskprint	Total Hashes	Total Files
AdvKeylogger-WinXP	4,716	23
Chrome28-W7x32	686,986	669
Chrome28-W7x64	670,051	499
Chrome28-WinXP	1,035,098	624
eraser-W7x32	69,984	24
Firefox19-W7x32	103,341	132
Firefox19-W7x64	106,270	146
Firefox19-WinXP	96,377	115
HxD171-W7x32	4,774	12
InvSecrets21-WinXP	6,689	19
OfficePro2003-W7x32	1,090,216	3,800
OfficePro2003-W7x64	1,077,126	3,804
OfficePro2003-WinXP	656,354	2,801
Python264-WinXP	86,287	2,355
Safari157-W7x32	316,224	907
Safari157-W7x64	569,645	1,504
Safari157-WinXP	343,824	918
sdelete-W7x32	642	5
sdelete-W7x64	642	4
Thunderbird2-WinXP	68,102	172
TrueCrypt63-WinXP	24,520	16
UPX-W7x32	1,796	19
UPX-W7x64	1,813	19
Winrar5beta-W7x32	9,196	41
Winrar5beta-W7x64	18,328	81
Winzip17pro-W7x32	240,229	149
Winzip17pro-W7x64	262,854	153
Wireshark-W7x32	171,515	617
Wireshark-W7x64	209,666	611
TOTALS	7,933,265	20,239

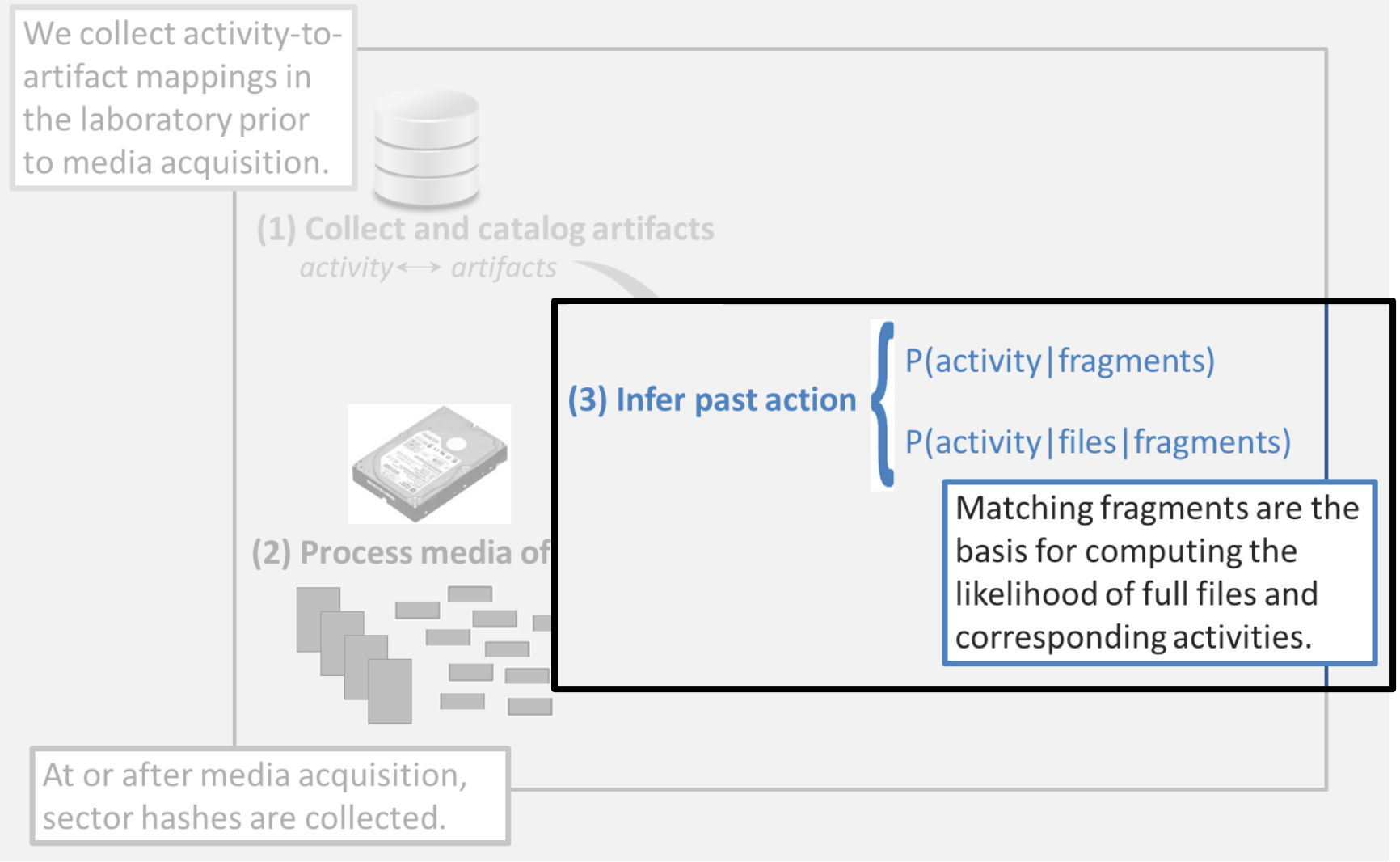
Step 2: Process media of interest



We hash the sectors on media of interest

- md5deep
- sector-aligned piecewise hashing
- 512-byte sectors
- Find matches using hashdb functionality
  - Matches with path, file, and diskprint ID

Step 3: Infer past action



Sector hits are weighted by catalog frequency

- Sector hits:
  - Original: sectors_found/ sectors_total
  - Weighted:

$$\text{weighted sector \%} = \left(\sum_{S=1}^{\text{num\_sec\_matches}} 1 / \text{freq}_S \right) / \text{sectors\_total}_{DP}$$

- Example:
  - Original: (1 + 1 + 1)/10 = 30%
  - Weighted: (1/1 + 1/4 + 1/2)/10 = 17.5%

File hits are weighted by % of file matched

- File hits:
  - Original: files_found/ files_total
  - Weighted:

$$\text{weighted file \%} = \left(\sum_{F=1}^{\text{num\_file\_matches}} \frac{\text{matched\_sectors}_F}{\text{total\_sectors}_F} \right) / \text{files\_total}_{DP}$$

- Example:
  - Original: (1 + 1)/5 = 40%
  - Weighted: (3/5 + 1/10)/5 = 14%

Sample output: Chrome Win7x64

diskprintName	sectors found	sectors total	sector%	w_sector%	files found	files total	file%	w_file%
Chrome28-W7x64	66795	670051	9.97%	3.63%	153	499	30.66%	21.46%
Chrome28-WinXP	40831	1035098	3.94%	1.16%	208	624	33.33%	21.10%
Chrome28-W7x32	66795	686986	9.72%	3.54%	152	669	22.72%	16.26%
Winzip17pro-W7x32	2186	240229	0.91%	0.46%	41	149	27.52%	3.63%
Winzip17pro-W7x64	2162	262854	0.82%	0.41%	42	153	27.45%	3.53%
Firefox19-W7x32	4183	103341	4.05%	0.59%	18	132	13.64%	2.44%
Firefox19-WinXP	4183	96377	4.34%	0.63%	17	115	14.78%	2.40%
Firefox19-W7x64	4184	106270	3.94%	0.57%	18	146	12.33%	2.37%
Thunderbird2-WinXP	17	68102	0.02%	0.01%	6	172	3.49%	1.09%
Winrar5beta-W7x64	9	18328	0.05%	0.01%	7	81	8.64%	0.38%
Winrar5beta-W7x32	9	9196	0.10%	0.02%	7	41	17.07%	0.38%
Safari157-WinXP	573	343824	0.17%	0.02%	31	918	3.38%	0.32%
Safari157-W7x32	573	316224	0.18%	0.02%	31	907	3.42%	0.30%
Safari157-W7x64	575	569645	0.10%	0.01%	35	1504	2.33%	0.24%
sdelete-W7x64	1	642	0.16%	0.04%	2	4	50.00%	0.17%
Wireshark-W7x32	51	171515	0.03%	0.01%	10	617	1.62%	0.16%
sdelete-W7x32	1	642	0.16%	0.04%	2	5	40.00%	0.14%
OfficePro2003-WinXP	1014	656354	0.15%	0.02%	33	2801	1.18%	0.13%
OfficePro2003-W7x32	1014	1090216	0.09%	0.01%	33	3800	0.87%	0.11%
OfficePro2003-W7x64	1014	1077126	0.09%	0.01%	33	3804	0.87%	0.08%
Wireshark-W7x64	11	209666	0.01%	0.00%	5	611	0.82%	0.02%
eraser-W7x32	21	69984	0.03%	0.02%	2	24	8.33%	0.01%
TrueCrypt63-WinXP	1	24520	0.00%	0.00%	1	16	6.25%	0.01%
Python264-WinXP	23	86287	0.03%	0.01%	6	2355	0.25%	0.00%
AdvKeylogger-WinXP	0	4716	0.00%	0.00%	0	23	0.00%	0.00%
InvSecrets21-WinXP	0	6689	0.00%	0.00%	0	19	0.00%	0.00%
UPX-W7x32	0	1796	0.00%	0.00%	0	19	0.00%	0.00%
HxD171-W7x32	0	4774	0.00%	0.00%	0	12	0.00%	0.00%
UPX-W7x64	0	1813	0.00%	0.00%	0	19	0.00%	0.00%

We created and processed eight test images (five single-app and three multi-app)

7 applications:

- Adv Keylogger
- **Chrome**
- Eraser
- **Firefox**
- HxD hex editor
- Invisible Secrets
- MS Office
- Python
- **Safari**
- **Sdelete**
- Thunderbird
- TrueCrypt
- **UPX**
- **WinRar**
- **WinZip**
- **Wireshark**

1 operating system:

- Windows XP (32 bit)
- Windows 7-32bit
- **Windows 7-64bit**

4-6 actions:

- Install
- App Activity:
  - Open
  - Use
  - Close
- Uninstall
- Reboot

Data set:

- 11 image sequences
  - 8 single and 3 multi-app
- 55 snapshots
  - 64 GB disks
- 55 hash sets
  - ~125M hashes each

	WinXP	Win7x32	Win7x64
Adv Keylogger	✓		
Chrome	✓	✓	✓
Eraser		✓	
Firefox	✓	✓	✓
HxD hex editor		✓	
Invisible Secrets	✓		
MS Office	✓	✓	✓
Python	✓		
Safari	✓	✓	✓
Sdelete		✓	✓
Thunderbird	✓		
TrueCrypt	✓		
UPX		✓	✓
WinRar		✓	✓
WinZip		✓	✓
Wireshark		✓	✓

Single application test case results

Source Image: Chrome28-W7x64		
diskprintName	**w_sector%**	**w_file%**
Chrome28-W7x64	3.63%	21.46%
Chrome28-WinXP	1.16%	21.10%
Chrome28-W7x32	3.54%	16.26%
Winzip17pro-W7x32	0.46%	3.63%
Winzip17pro-W7x64	0.41%	3.53%
...	...	...
Source Image: Winrar5beta-W7x64		
diskprintName	**w_sector%**	**w_file%**
Winrar5beta-W7x32	8.39%	56.18%
Winrar5beta-W7x64	4.21%	32.80%
Winzip17pro-W7x32	0.44%	3.53%
Winzip17pro-W7x64	0.41%	3.46%
sdelete-W7x32	0.04%	0.14%
...	...	...
Source Image: sdelete-W7x64		
diskprintName	**w_sector%**	**w_file%**
sdelete-W7x64	7.75%	33.95%
sdelete-W7x32	7.75%	27.16%
Winzip17pro-W7x32	0.44%	3.52%
Winzip17pro-W7x64	0.41%	3.45%
Firefox19-W7x64	0.01%	1.67%
...	...	...

Source Image: UPX-W7x64		
diskprintName	**w_sector%**	**w_file%**
UPX-W7x32	2.97%	52.16%
UPX-W7x64	2.94%	52.16%
Winzip17pro-W7x32	0.44%	3.52%
Winzip17pro-W7x64	0.41%	3.45%
Firefox19-W7x64	0.01%	1.69%
...	...	...
Source Image: Firefox19-W7x64		
diskprintName	**w_sector%**	**w_file%**
Firefox19-WinXP	6.88%	57.32%
Firefox19-W7x32	6.42%	51.52%
Firefox19-W7x64	6.26%	47.25%
Winzip17pro-W7x32	0.44%	3.57%
Winzip17pro-W7x64	0.41%	3.48%
...	...	...

Multiple application test case results

Source Image: Chrome & Firefox	
diskprintName	**w_file%**
Firefox19-WinXP	57.21%
Firefox19-W7x32	52.04%
Firefox19-W7x64	47.33%
Chrome28-W7x64	20.45%
Chrome28-WinXP	20.37%
Chrome28-W7x32	15.58%
Winzip17pro-W7x32	3.64%
Winzip17pro-W7x64	3.53%
Thunderbird2-WinXP	1.68%
Winrar5beta-W7x64	0.42%
...	...

Source Image: WinRAR & WinZip	
diskprintName	**w_file%**
Winzip17pro-W7x64	35.60%
Winzip17pro-W7x32	34.88%
Winrar5beta-W7x32	9.97%
Winrar5beta-W7x64	9.29%
Firefox19-WinXP	2.66%
Firefox19-W7x64	2.60%
Firefox19-W7x32	2.23%
Thunderbird2-WinXP	1.49%
sdelete-W7x64	0.17%
Wireshark-W7x32	0.16%
...	...

Source Image: Firefox, Chrome, & Safari	
diskprintName	**w_file%**
Safari157-W7x32	94.32%
Safari157-WinXP	92.80%
Safari157-W7x64	57.12%
Firefox19-WinXP	46.57%
Firefox19-W7x32	42.83%
Firefox19-W7x64	37.73%
Chrome28-WinXP	22.10%
Chrome28-W7x64	12.88%
Chrome28-W7x32	9.84%
Winzip17pro-W7x32	3.62%
...	...

M57 test case results

Charlie		Jo		Pat		Terry	
diskprintName	w_file%	diskprintName	w_file%	diskprintName	w_file%	diskprintName	w_file%
Python264-WinXP	98.98%	Python264-WinXP	98.83%	Python264-WinXP	98.91%	Python264-WinXP	85.52%
InvSecrets21-WinXP	63.16%	TrueCrypt63-WinXP	50.00%	Thunderbird2-WinXP	24.94%	Thunderbird2-WinXP	27.81%
Thunderbird2-WinXP	61.00%	Thunderbird2-WinXP	24.73%	AdvKeylogger-WinXP	21.97%	Winzip17pro-W7x64	10.37%
Safari157-W7x32	10.25%	Safari157-W7x32	11.35%	HxD171-W7x32	8.39%	Winzip17pro-W7x32	10.05%
Safari157-WinXP	10.16%	Safari157-WinXP	11.26%	Firefox19-WinXP	3.17%	HxD171-W7x32	8.37%
Safari157-W7x64	6.69%	Safari157-W7x64	7.37%	Firefox19-W7x64	2.93%	Safari157-W7x32	5.46%
Firefox19-WinXP	3.26%	Firefox19-WinXP	3.24%	Firefox19-W7x32	2.78%	Safari157-WinXP	5.35%
Firefox19-W7x32	2.77%	Firefox19-W7x32	2.74%	Winzip17pro-W7x64	2.03%	Chrome28-WinXP	4.83%
Firefox19-W7x64	2.50%	Firefox19-W7x64	2.62%	Chrome28-WinXP	1.64%	Chrome28-W7x64	4.81%
Chrome28-WinXP	2.11%	Chrome28-WinXP	2.15%	Chrome28-W7x64	1.63%	Firefox19-WinXP	3.59%
Winzip17pro-W7x64	2.08%	Chrome28-W7x64	2.03%	Winzip17pro-W7x32	1.50%	Chrome28-W7x32	3.59%
Chrome28-W7x64	2.02%	Chrome28-W7x32	1.52%	Chrome28-W7x32	1.22%	Firefox19-W7x64	3.56%
Chrome28-W7x32	1.52%	sdelete-W7x64	1.35%	TrueCrypt63-WinXP	1.22%	Firefox19-W7x32	3.55%
Winzip17pro-W7x32	1.51%	Winzip17pro-W7x64	1.26%	Winrar5beta-W7x64	0.85%	Safari157-W7x64	3.47%
sdelete-W7x64	1.35%	sdelete-W7x32	1.08%	Winrar5beta-W7x32	0.84%	Winrar5beta-W7x64	2.21%
sdelete-W7x32	1.08%	Winrar5beta-W7x64	0.95%	Safari157-WinXP	0.62%	Winrar5beta-W7x32	2.19%
TrueCrypt63-WinXP	0.73%	Winrar5beta-W7x32	0.94%	Safari157-W7x32	0.54%	TrueCrypt63-WinXP	0.97%
Winrar5beta-W7x32	0.64%	Winzip17pro-W7x32	0.72%	OfficePro2k3-WinXP	0.47%	OfficePro2k3-W7x32	0.39%
Winrar5beta-W7x64	0.64%	OfficePro2k3-WinXP	0.43%	OfficePro2k3-W7x32	0.45%	OfficePro2k3-WinXP	0.35%
OfficePro2k3-WinXP	0.37%	OfficePro2k3-W7x32	0.41%	OfficePro2k3-W7x64	0.42%	OfficePro2k3-W7x64	0.35%
OfficePro2k3-W7x32	0.32%	OfficePro2k3-W7x64	0.37%	Safari157-W7x64	0.39%	Wireshark-W7x32	0.09%
OfficePro2k3-W7x64	0.31%	Wireshark-W7x32	0.07%	Wireshark-W7x32	0.10%	eraser-W7x32	0.05%
Wireshark-W7x32	0.06%	HxD171-W7x32	0.04%	Wireshark-W7x64	0.02%	Wireshark-W7x64	0.05%
eraser-W7x32	0.01%	eraser-W7x32	0.02%	eraser-W7x32	0.02%	AdvKeylogger-WinXP	0.03%
AdvKeylogger-WinXP	0.01%	Wireshark-W7x64	0.02%	InvSecrets21-WinXP	0.00%	InvSecrets21-WinXP	0.00%
Wireshark-W7x64	0.00%	AdvKeylogger-WinXP	0.01%	UPX-W7x32	0.00%	UPX-W7x32	0.00%
UPX-W7x32	0.00%	InvSecrets21-WinXP	0.00%	sdelete-W7x32	0.00%	sdelete-W7x32	0.00%
HxD171-W7x32	0.00%	UPX-W7x32	0.00%	UPX-W7x64	0.00%	UPX-W7x64	0.00%

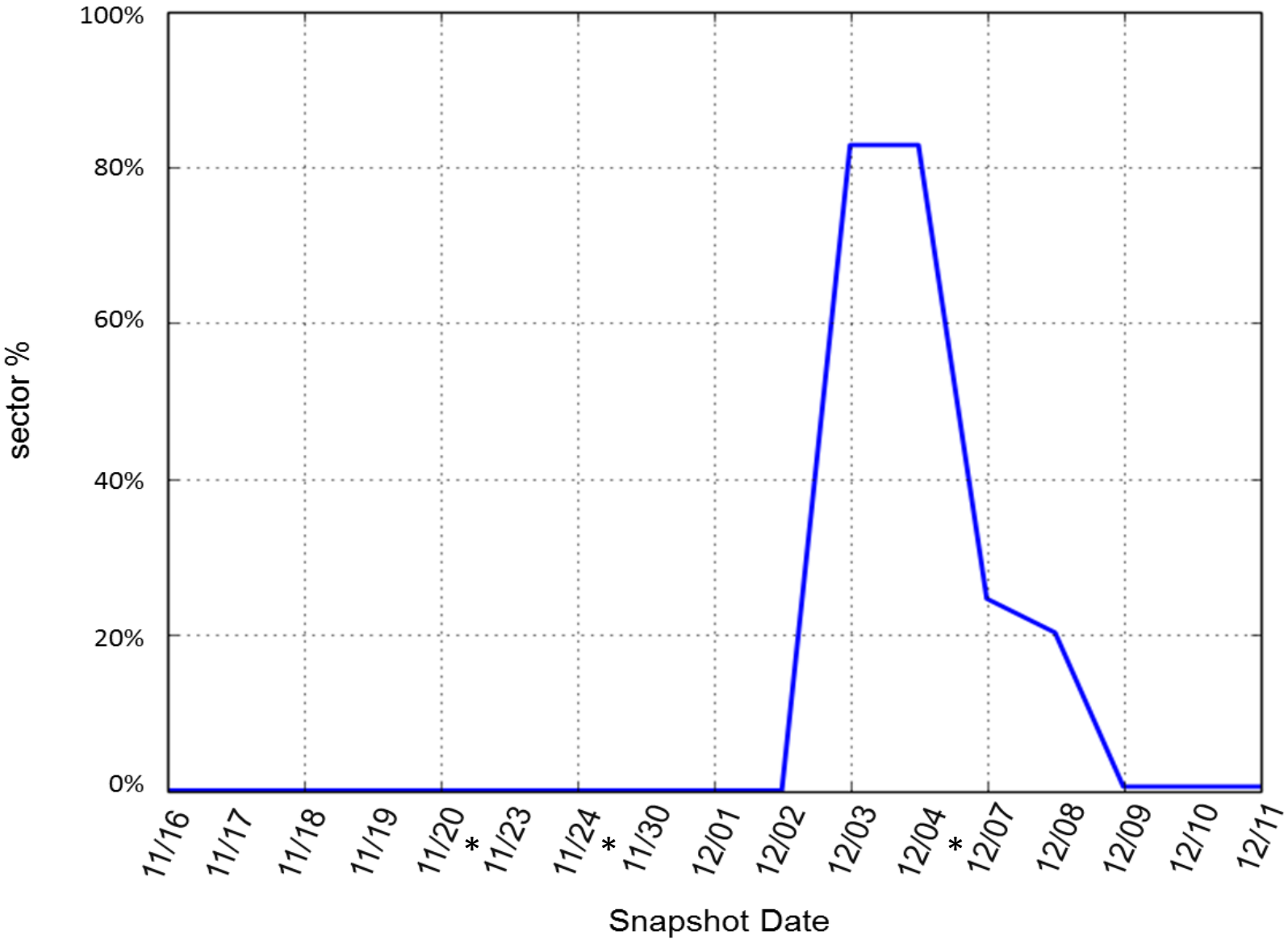
Legend
Confirmed
Suspected
Unconfirmed

M57 test case results

Charlie		Jo		Pat		Terry	
diskprintName	w_file%	diskprintName	w_file%	diskprintName	w_file%	diskprintName	w_file%
Python264-WinXP	98.98%	Python264-WinXP	98.83%	Python264-WinXP	98.91%	Python264-WinXP	85.52%
InvSecrets21-WinXP	63.16%	TrueCrypt63-WinXP	50.00%	Thunderbird2-WinXP	24.94%	Thunderbird2-WinXP	27.81%
Thunderbird2-WinXP	61.00%	Thunderbird2-WinXP	24.73%	AdvKeylogger-WinXP	21.97%	Winzip17pro-W7x64	10.37%
Safari157-W7x32	10.25%	Safari157-W7x32	11.35%	HxD171-W7x32	8.39%	Winzip17pro-W7x32	10.05%
Safari157-WinXP	10.16%	Safari157-WinXP	11.26%	Firefox19-WinXP	3.17%	HxD171-W7x32	8.37%
Safari157-W7x64	6.69%	Safari157-W7x64	7.37%	Firefox19-W7x64	2.93%	Safari157-W7x32	5.46%
Firefox19-WinXP	3.26%	Firefox19-WinXP	3.24%	Firefox19-W7x32	2.78%	Safari157-WinXP	5.35%
Firefox19-W7x32	2.77%	Firefox19-W7x32	2.74%	Winzip17pro-W7x64	2.03%	Chrome28-WinXP	4.83%
Firefox19-W7x64	2.50%	Firefox19-W7x64	2.62%	Chrome28-WinXP	1.64%	Chrome28-W7x64	4.81%
Chrome28-WinXP	2.11%	Chrome28-WinXP	2.15%	Chrome28-W7x64	1.63%	Firefox19-WinXP	3.59%
Winzip17pro-W7x64	2.08%	Chrome28-W7x64	2.03%	Winzip17pro-W7x32	1.50%	Chrome28-W7x32	3.59%
Chrome28-W7x64	2.02%	Chrome28-W7x32	1.52%	Chrome28-W7x32	1.22%	Firefox19-W7x64	3.56%
Chrome28-W7x32	1.52%	sdelete-W7x64	1.35%	TrueCrypt63-WinXP	1.22%	Firefox19-W7x32	3.55%
Winzip17pro-W7x32	1.51%	Winzip17pro-W7x64	1.26%	Winrar5beta-W7x64	0.85%	Safari157-W7x64	3.47%
sdelete-W7x64	1.35%	sdelete-W7x32	1.08%	Winrar5beta-W7x32	0.84%	Winrar5beta-W7x64	2.21%
sdelete-W7x32	1.08%	Winrar5beta-W7x64	0.95%	Safari157-WinXP	0.62%	Winrar5beta-W7x32	2.19%
TrueCrypt63-WinXP	0.73%	Winrar5beta-W7x32	0.94%	Safari157-W7x32	0.54%	TrueCrypt63-WinXP	0.97%
Winrar5beta-W7x32	0.64%	Winzip17pro-W7x32	0.72%	OfficePro2k3-WinXP	0.47%	OfficePro2k3-W7x32	0.39%
Winrar5beta-W7x64	0.64%	OfficePro2k3-WinXP	0.43%	OfficePro2k3-W7x32	0.45%	OfficePro2k3-WinXP	0.35%
OfficePro2k3-WinXP	0.37%	OfficePro2k3-W7x32	0.41%	OfficePro2k3-W7x64	0.42%	OfficePro2k3-W7x64	0.35%
OfficePro2k3-W7x32	0.32%	OfficePro2k3-W7x64	0.37%	Safari157-W7x64	0.39%	Wireshark-W7x32	0.09%
OfficePro2k3-W7x64	0.31%	Wireshark-W7x32	0.07%	Wireshark-W7x32	0.10%	eraser-W7x32	0.05%
Wireshark-W7x32	0.06%	HxD171-W7x32	0.04%	Wireshark-W7x64	0.02%	Wireshark-W7x64	0.05%
eraser-W7x32	0.01%	eraser-W7x32	0.02%	eraser-W7x32	0.02%	AdvKeylogger-WinXP	0.03%
AdvKeylogger-WinXP	0.01%	Wireshark-W7x64	0.02%	InvSecrets21-WinXP	0.00%	InvSecrets21-WinXP	0.00%
Wireshark-W7x64	0.00%	AdvKeylogger-WinXP	0.01%	UPX-W7x32	0.00%	UPX-W7x32	0.00%
UPX-W7x32	0.00%	InvSecrets21-WinXP	0.00%	sdelete-W7x32	0.00%	sdelete-W7x32	0.00%
HxD171-W7x32	0.00%	UPX-W7x32	0.00%	UPX-W7x64	0.00%	UPX-W7x64	0.00%

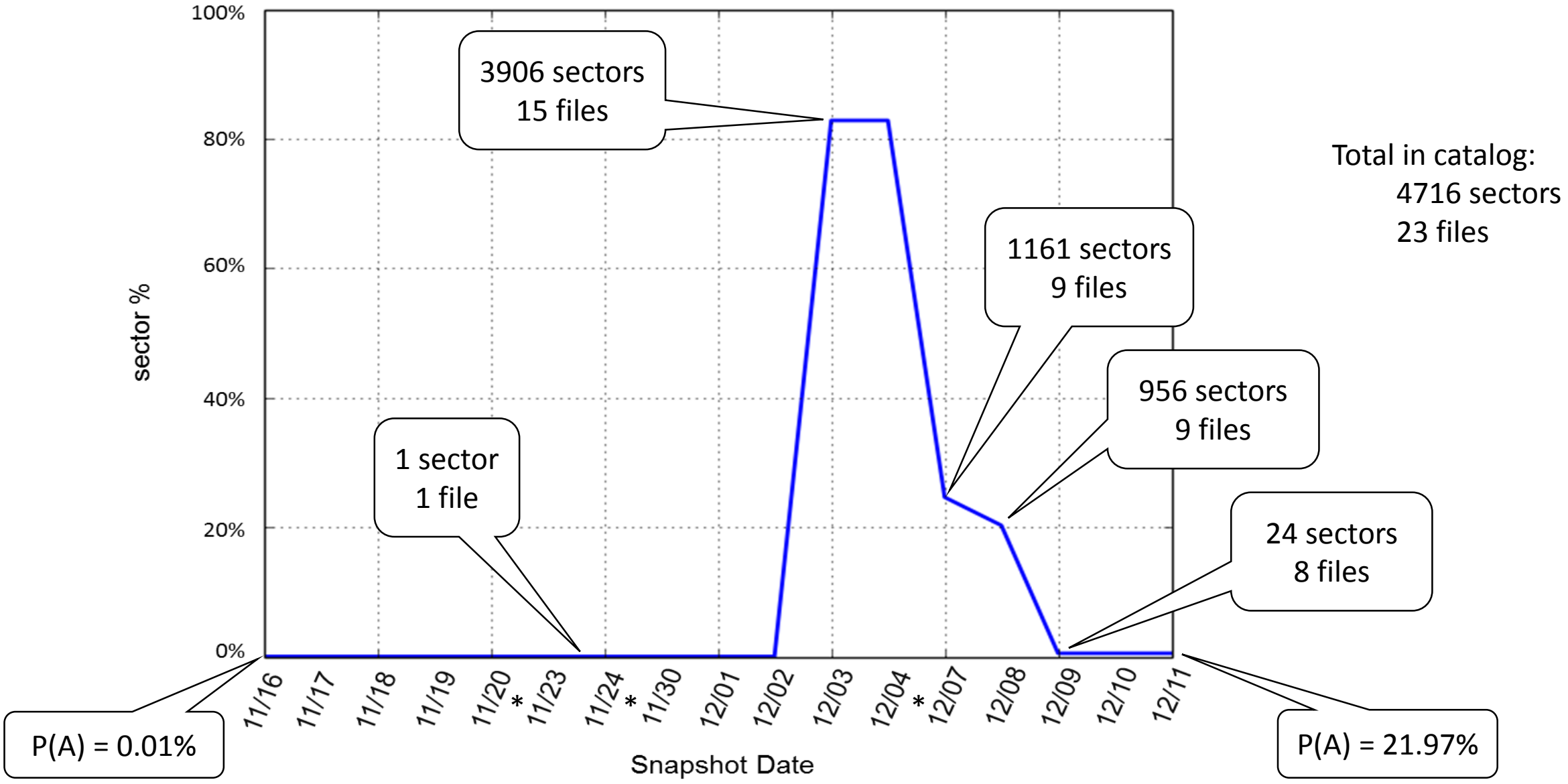
Legend
Confirmed
Suspected
Unconfirmed

Advanced Keylogger (Pat) artifact persistence



Total in catalog:
4716 sectors
23 files

Advanced Keylogger (Pat) artifact persistence



Total in catalog:
4716 sectors
23 files

Current and future work...

- Autopsy Plugin
- Extensions of this work:
  - enhance computation
  - noise reduction at collection
  - instrumented collection
  - sector differencing
- Artifact persistence model
- Mobile
- Malware
- Memory

Current and future work...

- Autopsy Plugin
- Extensions of this work:
  - enhance computation
  - noise reduction at collection
  - instrumented collection
  - sector differencing
- Artifact persistence model
- Mobile
- Malware
- Memory

Questions?