THE UNCERTAIN RELATIONSHIP BETWEEN RESIDUAL DIGITAL FRAGMENTS, FILES, AND COMPUTER ACTIVITY

Jim Jones, PhD

Associate Professor

Digital Forensics and Cyber Analysis

George Mason University

2017 International Symposium on Forensic Science Error Management

THE CHALLENGE: CAN WE IDENTIFY PAST APPLICATION USE AFTER UNINSTALLATION?

Malicious Windows user installs wireshark...

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

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

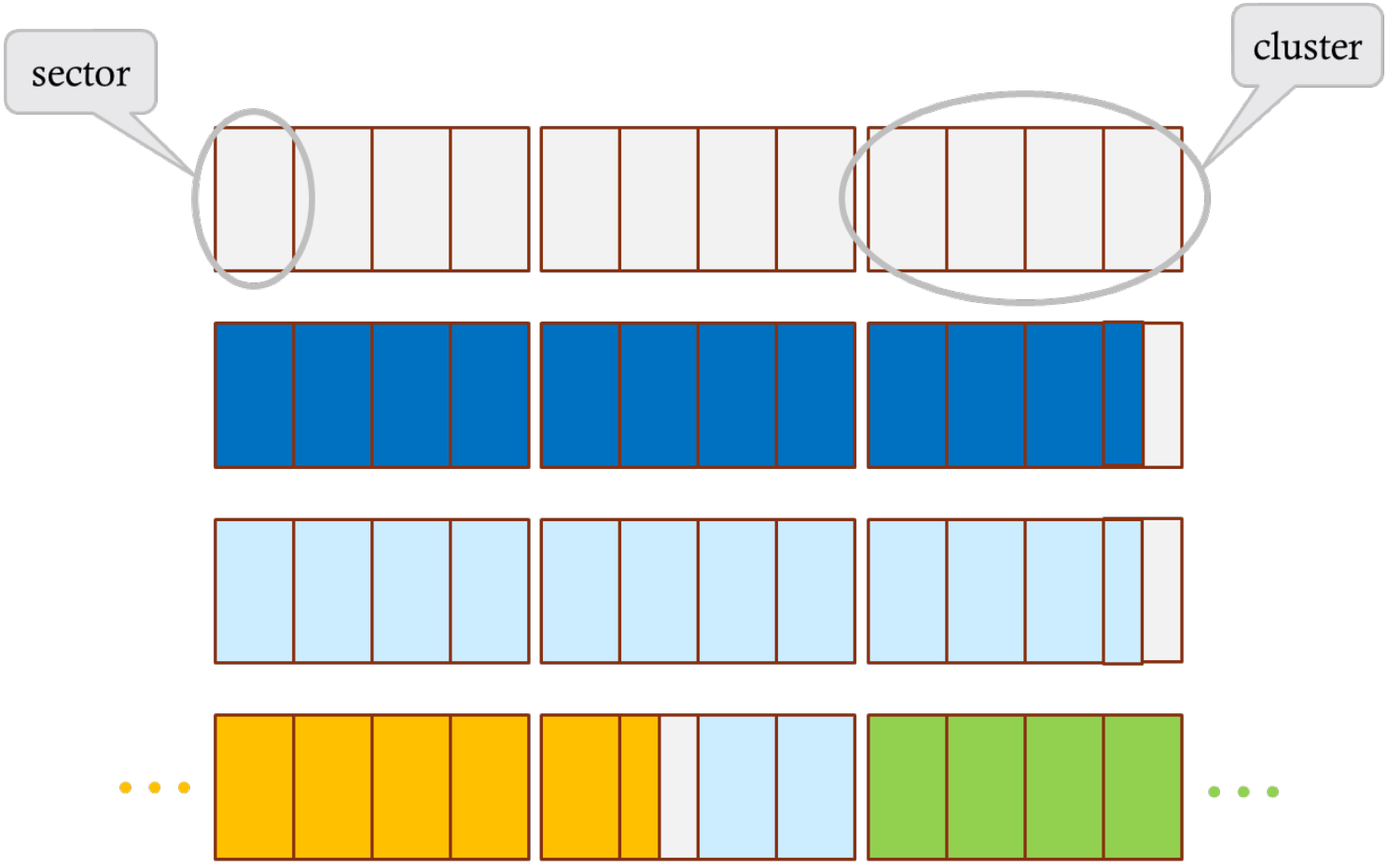
...uses account info to commit a crime such as impersonate another user or access sensitive information...

...law enforcement confiscates computer...

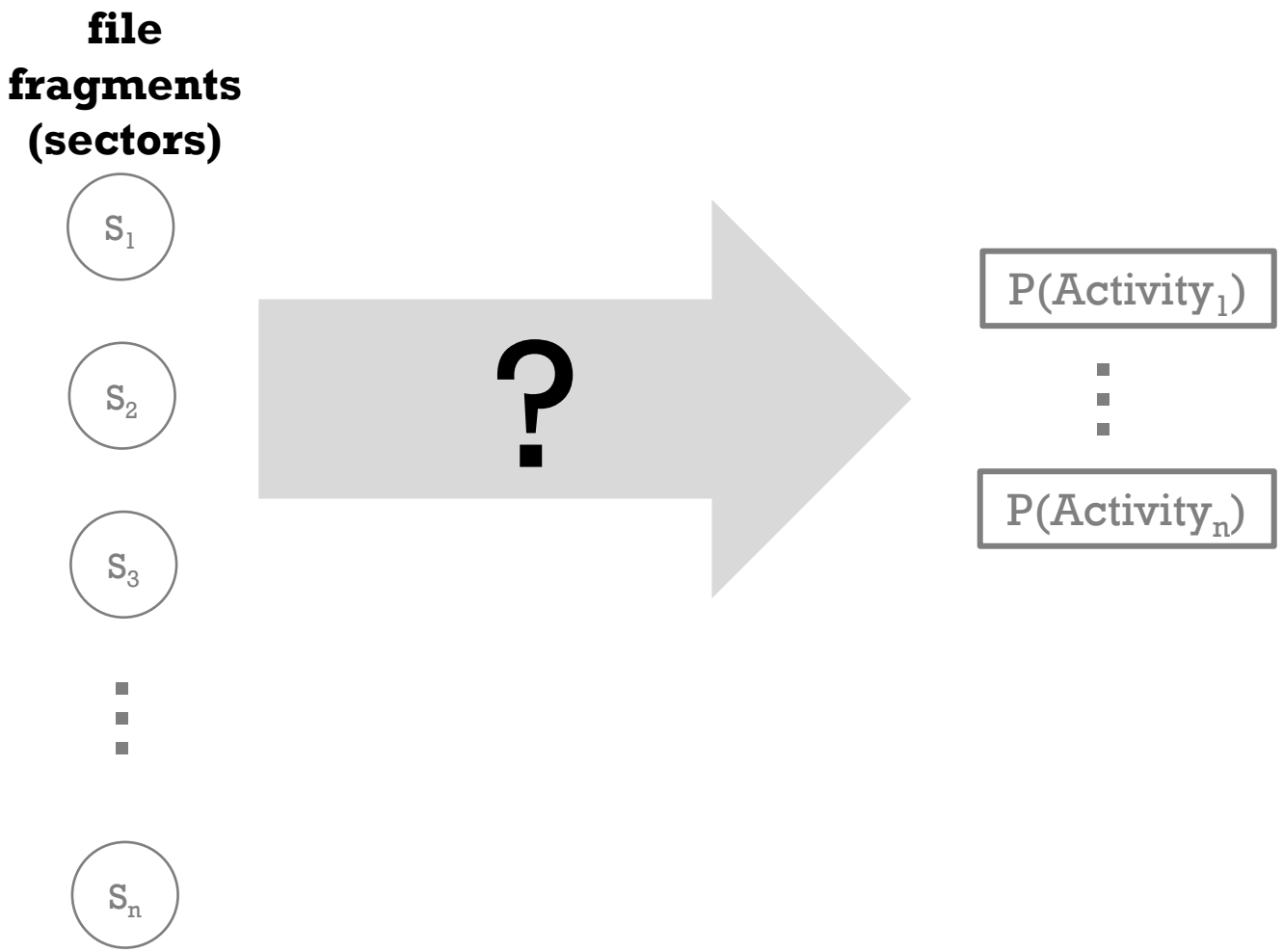
...can we show wireshark was used in the past?

Image credits:
http://school.discoveryeducation.com/clipart/clip/stk-fgr6.html
http://blog.deming.org/2014/10/the-target-is-irrelevant-without-a-method/
http://www.presentermedia.com/index.php?target=closeup&maincat=clipart&id=9771

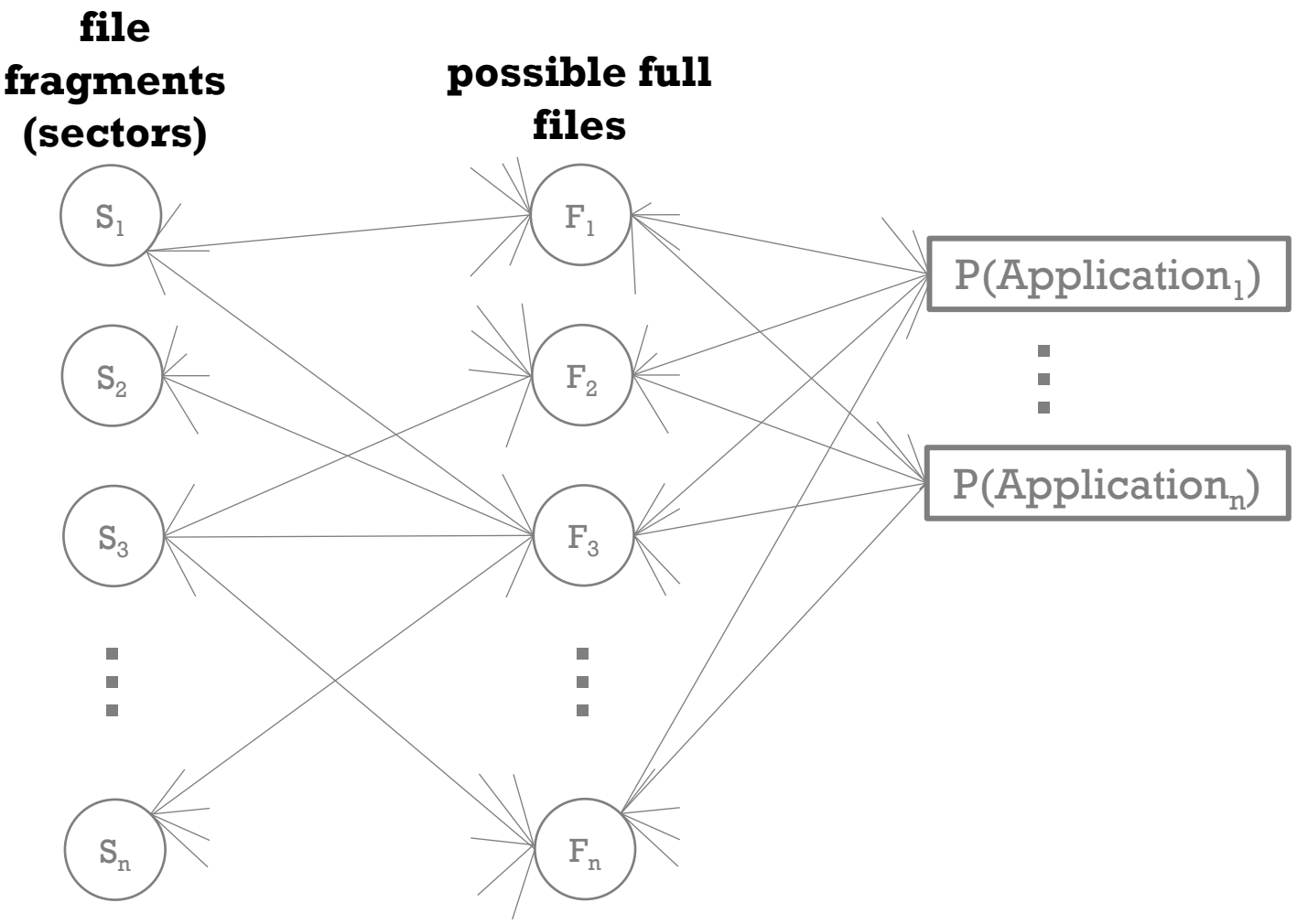
RESIDUAL FILE FRAGMENTS ARE CREATED AFTER A FILE IS DELETED



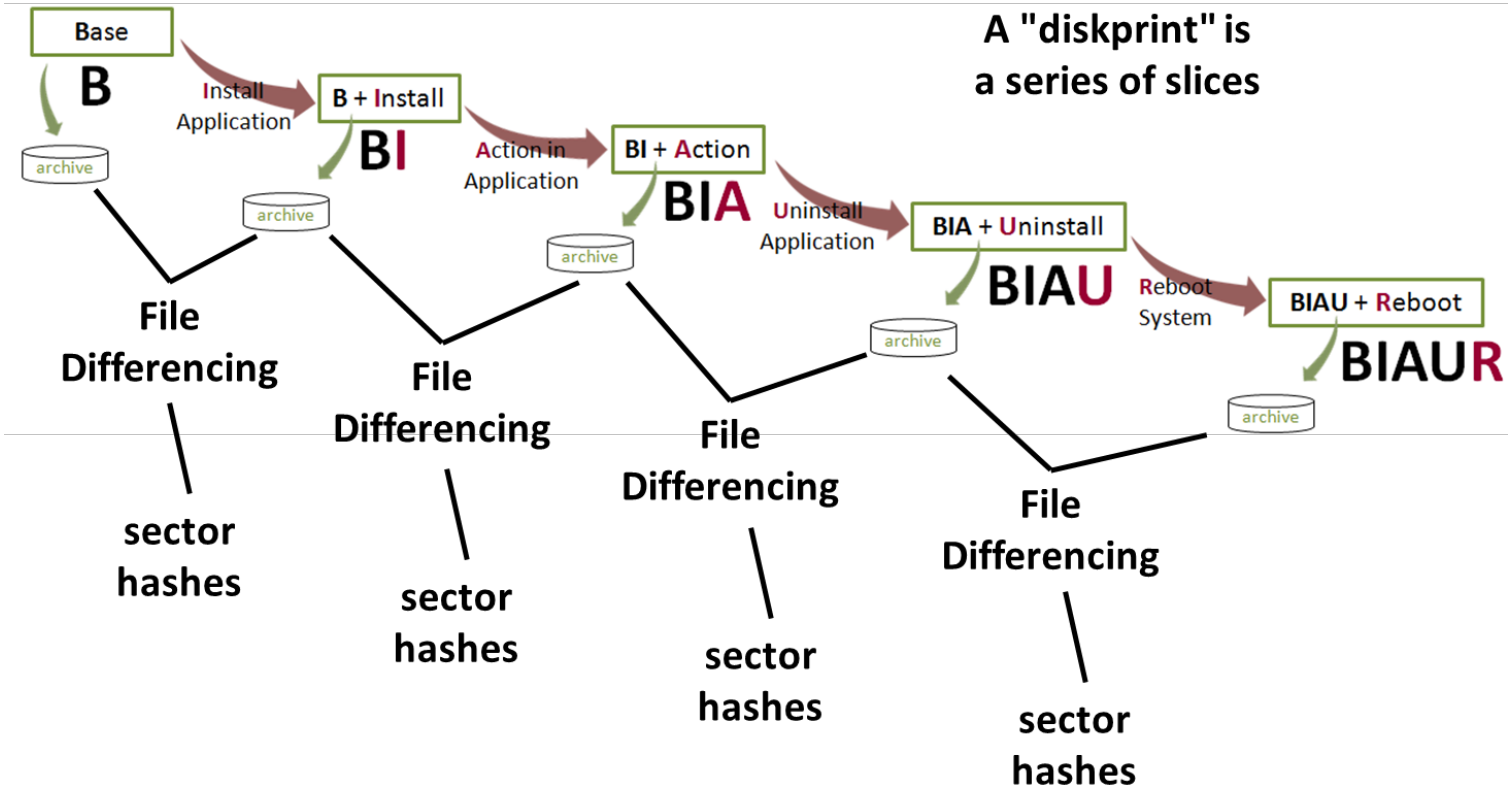
CAN WE INFER PAST APPLICATION USE FROM DELETED FILE FRAGMENTS?



APPROACH: REASON FROM FRAGMENTS TO FILES AND FILES TO APPLICATIONS



PRIOR KNOWLEDGE: APPLICATIONS TO FILES TO FRAGMENTS



FRAGMENTS TO FILES...

- 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%

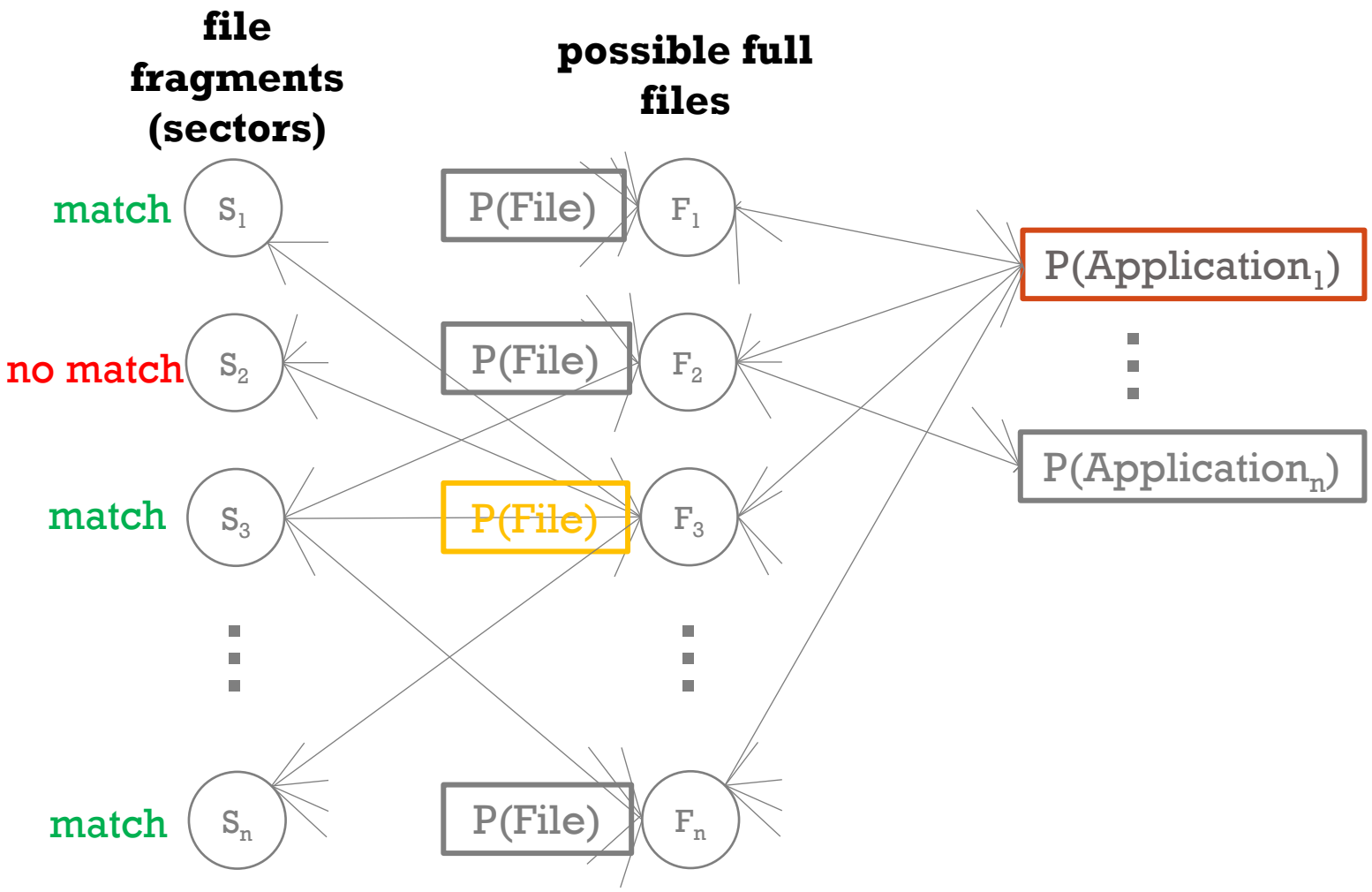
FILES TO APPLICATIONS…

- 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%

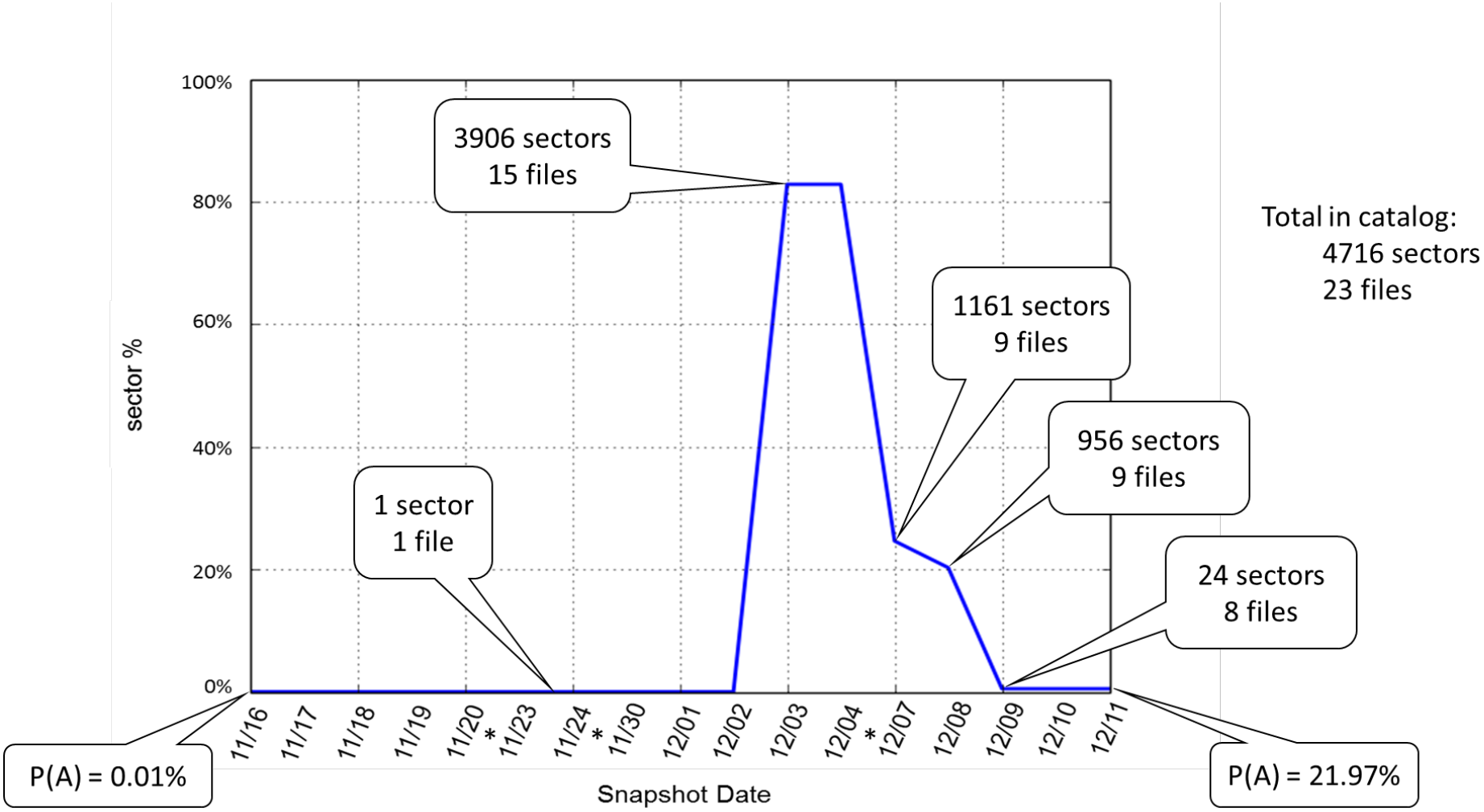
TBD: COMBINE BOTH WEIGHTED MEASURES



EXPERIMENTS: M57 PATENTS DATASET

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

ADVANCED KEYLOGGER EXAMPLE



Total in catalog:
4716 sectors
23 files

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%

CHALLENGES AND CURRENT WORK

- Combine weighted measures:
  - fragments to files to applications
- Weighting factors are incomplete:
  - relative location, file offset, match frequency, entropy, …
- We don't have the real "universe" of files:
  - NSRL, Google, Cloud providers, …
- Deleted file decay is inconsistent; depends (at least) on:
  - OS, filesystem, hardware, user and system activity, format, disk usage and fragmentation, file characteristics, …
  - time matters: can we use this to date the application uninstallation?
- Other scenarios:
  - mobile, malware, ICS, …

ANY QUESTIONS, IDEAS, OR SUGGESTIONS...?

Jim Jones, PhD
Associate Professor, ECE/DFCA
Nguyen Engineering Bldg., Room 3241
George Mason University, MS 2B5
Fairfax, VA 22030
(o) 703-993-5599
(c) 703-955-1033
(e) jjonesu@gmu.edu
(w) http://ece.gmu.edu/
(w) http://cfrs.gmu.edu/