2013 IEEE International Conference on
Technologies for Homeland Security
12-14 November 2013, Westin Hotel, Waltham, MA

IEEE
Advancing Technology for Humanity

DETECTING THREATENING INSIDERS WITH LIGHTWEIGHT MEDIA FORENSICS

Naval Postgraduate School &
The University of Texas at San Antonio

Dr. Simson Garfinkel (NPS) & Dr. Nicole Beebe (UTSA)

8am, Wednesday November 13th, 2013

NPS PRAESTANTIA PER SCIENTIAM 1909

UTSA

DEPARTMENT OF DEFENSE UNITED STATES OF AMERICA

ASD(R&E)

Team Profile

Naval Postgraduate School

- Simson L. Garfinkel
  Assoc. Prof
  Computer Science
  —*simsong@acm.org*
  —*+1.202.649.0029*

The University of Texas at San Antonio

- N. Beebe, Asst. Prof.
  Info Systems/Cyber Security
  —*Nicole.Beebe@utsa.edu*
  —*+1.210.269.5647*

The current approaches for finding hostile insiders are based on "signatures."

Sample signature to find a problem employee:

(CERT 2011)

- ***if the mail is from a departing insider***
- ***and the message was sent in last 30 days***
- ***and the recipient is not in organization's domain***
- ***and the total bytes summed by day is more than X,***

➜ ***send an alert to security operator***

These signatures are typically hand written.

—*Brittle*

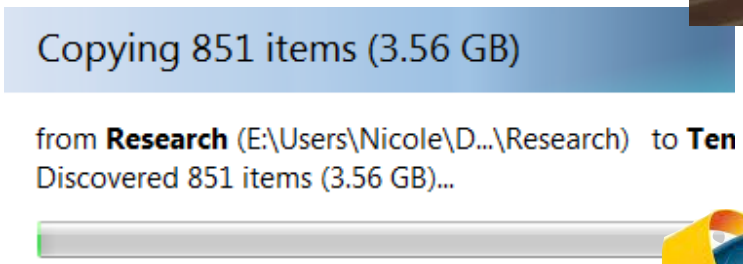
—*Don't scale*

—*Miss new patterns*

We propose a new approach for finding threatening insiders—storage profile anomalies.

Hypothesis 1:
Some insiders hoard before exfiltration

- Manning
- Snowden



We also want to detect other kinds of illegal employee activity.

Hypothesis 2:
Some illegal activity has storage indicators:

- Contraband software (hacking tools) and data
- Large amount of:
  - —*graphics*
  - —*PII; PHI; account numbers*
  - —*Encrypted data*
- Stolen documents

Illegal employee activity is:

- Bad for business
- Exploitation threat
- Fraud risk

CNN Justice

Pentagon reopening probe into employees allegedly tied to child porn

By Adam Levine, CNN
September 16, 2010 11:59 a.m. EDT

The Defense Department will review 264 cases of possible trafficking in child pornography.

(CNN) -- The Defense Department will reopen its investigation into employees who are alleged to have downloaded child pornography, a spokesman said Wednesday.

The Pentagon's Defense Criminal Investigative Service will review 264 cases, according to spokesman Gary Comerford. The department had stopped the reviews because of a lack of resources, he said.

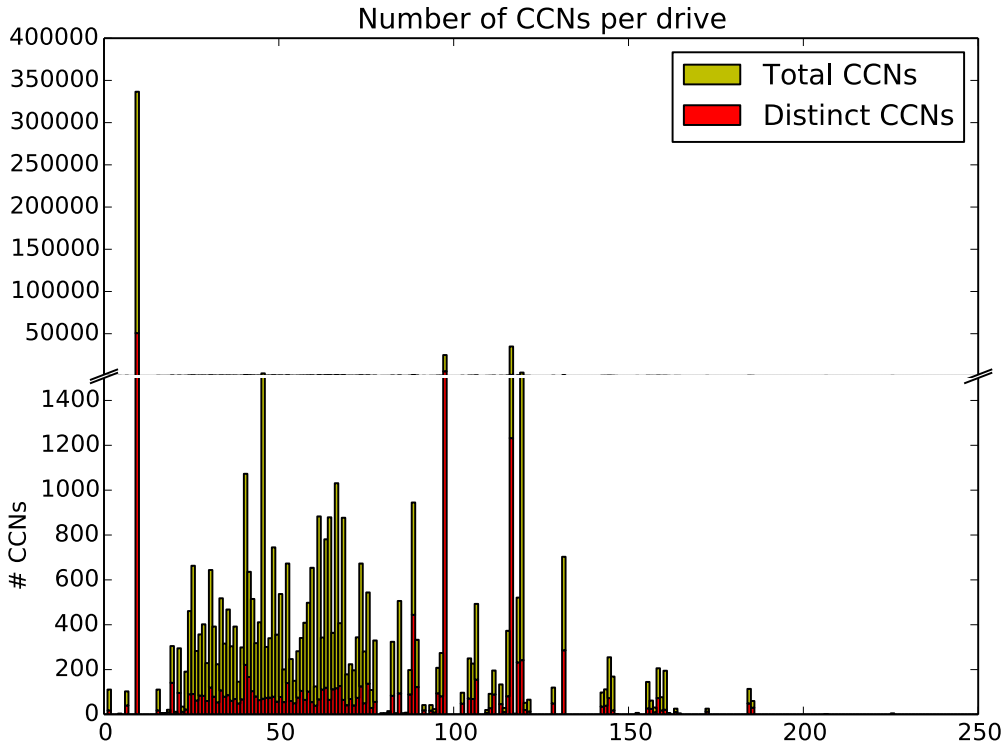
Our plan: look for storage devices that are different than their peers.

We build a “storage profile” from features:

- # of credit card numbers, phone #s; SSNs, DOBs, etc.
- % pictures; %video
- % Doc files; %PDFs;

“Different” relative to:

- User’s history
- User’s organization
- Others in role.

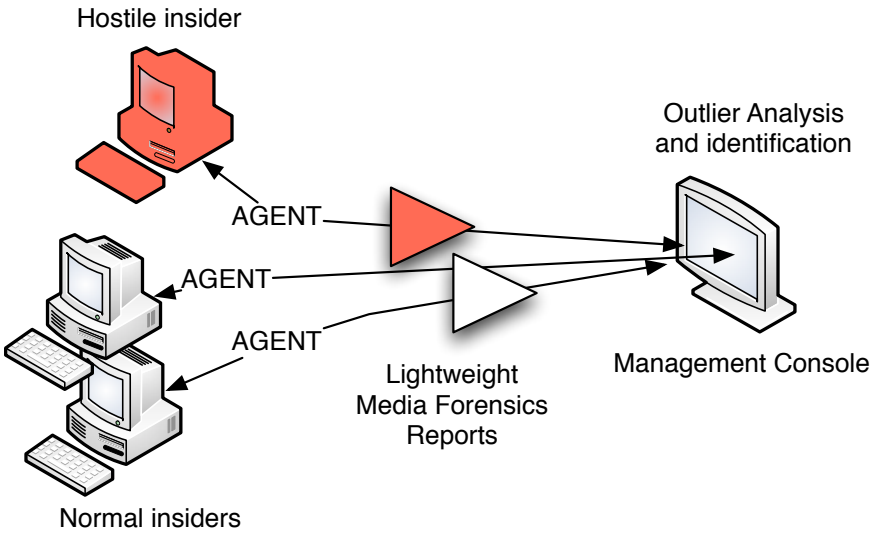


Garfinkel, S. and Shelat, A., "Remembrance of Data Passed: A Study of Disk Sanitization Practices," IEEE Security & Privacy, January/February 2003.

Our approach:
Collect “storage profiles” and look for outliers.

We profile storage on the hard drive/storage device:

- Allocated & “deleted” files; Unallocated space (file fragments)



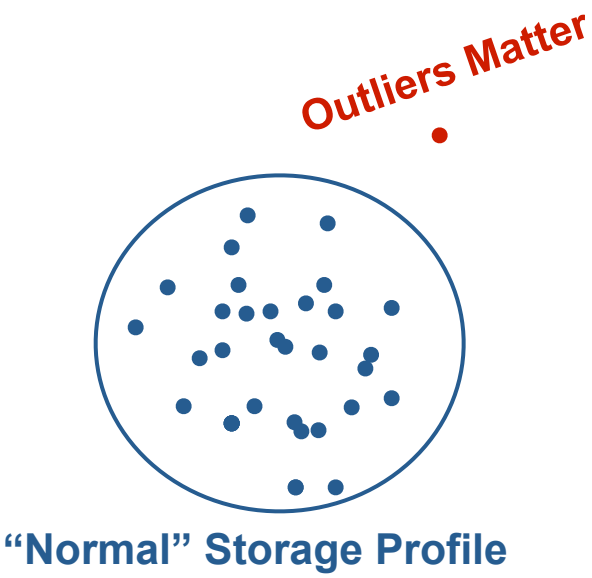
Statistical profile is collected:

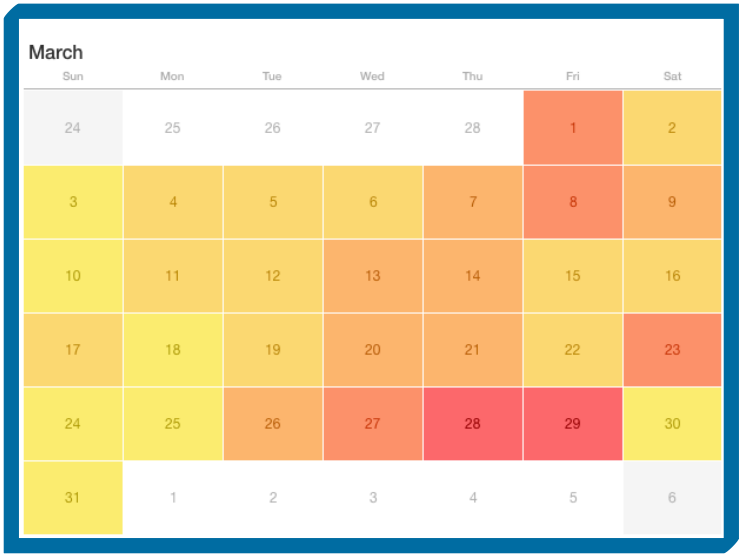
- Frequently, at “random” times
- Securely — by going to raw media
- Centrally — at management console

We cluster the storage profiles to find “outliers.”

What’s an outlier?

- Something that’s different from its peers
- Something different from its own history





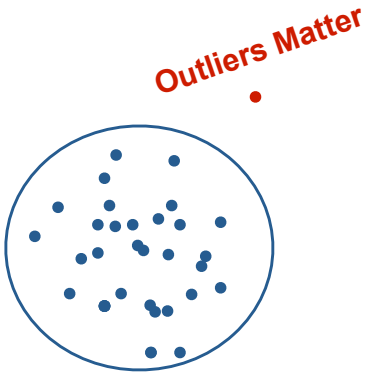
Outlier detection should have significant benefits:

- Not signature based
- Not reliant on access patterns
- Not reliant on policy definition, discovery, auditing

Design constraints:

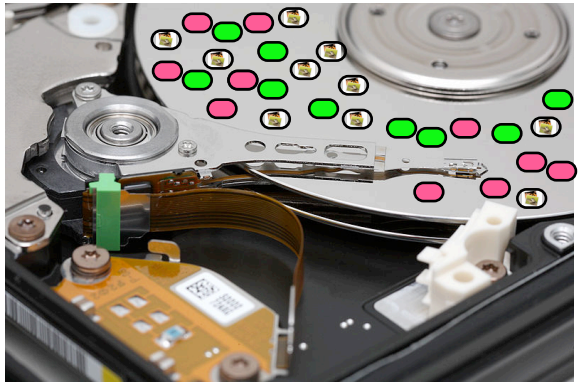
- Agent must be scalable and cannot interfer with operations
  - —*Desktop: background process, samples disk data*
  - —*Network load: small, aggregated data transfer*
  - —*Management console: scalable algorithms used*

- Must work with isolated systems
- Must be OS agnostic
- Must includes deleted data in collection/analysis



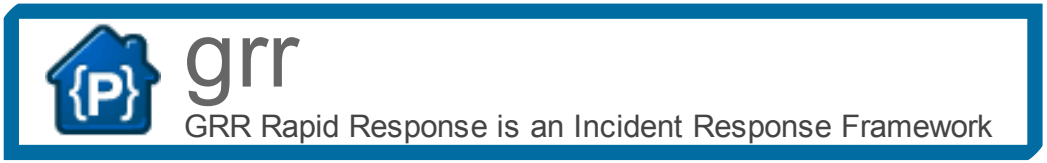
Our system has three parts:

1. Sample disk to collect desired data
   - bulk_extractor
     — *a lightweight media forensics tool*

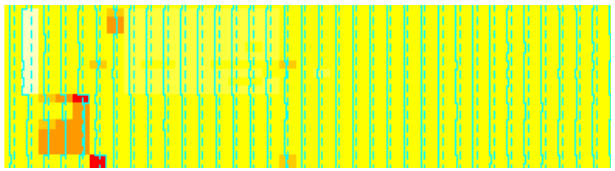


Garfinkel, Simson, Digital media triage with bulk data analysis and bulk_extractor. Computers and Security 32: 56-72 (2013)

2. Client-server, enterprise response framework
   - Google Rapid Response (GRR)

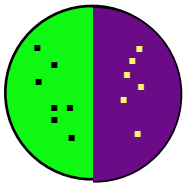


3. Anomaly detection agent
   - Univariate and multivariate outlier detection



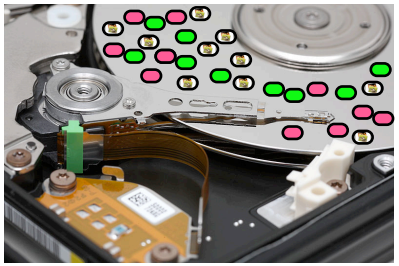
Random sampling is a great way to analyze data.

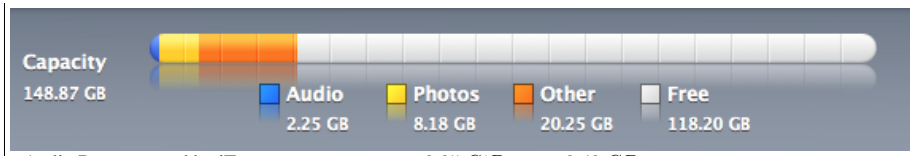
Simple random sampling can determine % free space



Garfinkel, Simson, Vassil Roussev, Alex Nelson and Douglas White, Using purpose-built functions and block hashes to enable small block and sub-file forensics, DFRWS 2010, Portland, OR

Data characterization can determine the *kind* of stored data

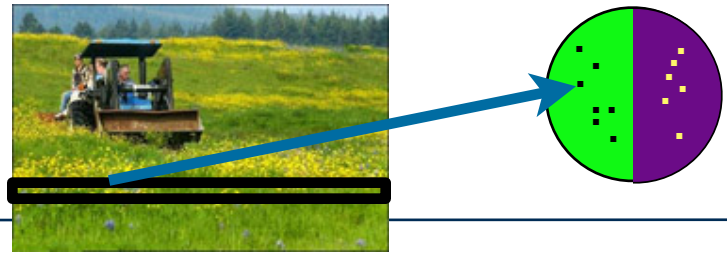




Audio Data reported by iTunes:	2.25 *GiB*	2.42 *GB*	
MP3 files reported by file system:		2.39 *GB*	
Estimated MP3 usage with random sampling :		2.49 *GB*	10,000 random samples
		2.71 *GB*	5,000 random samples

Sector hashing can identify *specific target files*

Young J., Foster, K., Garfinkel, S., and Fairbanks, K., Distinct sector hashes for target file detection, IEEE Computer, December 2012



It takes 3.5 hours to read a 1TB hard drive.

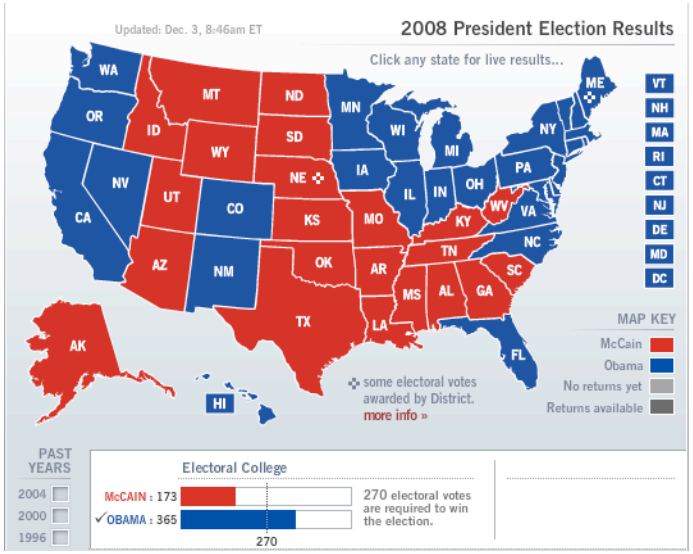
In 5 minutes you can read:

- 36 GB in one strip
- 100,000 randomly chosen 64KiB strips (assuming 3 msec/seek)

			
Minutes	208	5	5
Data	1 TB	36 GB	6.5 GB
# Seeks	1	1	100,000
% of data	100%	3.6%	0.65%

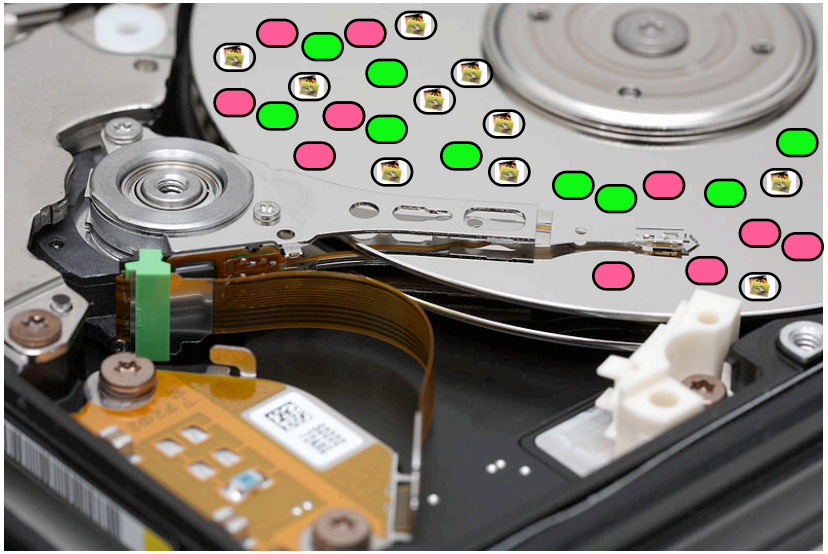
The statistics of a *randomly chosen sample* predict the *statistics of a population.*

US elections can be predicted by sampling thousands of households:



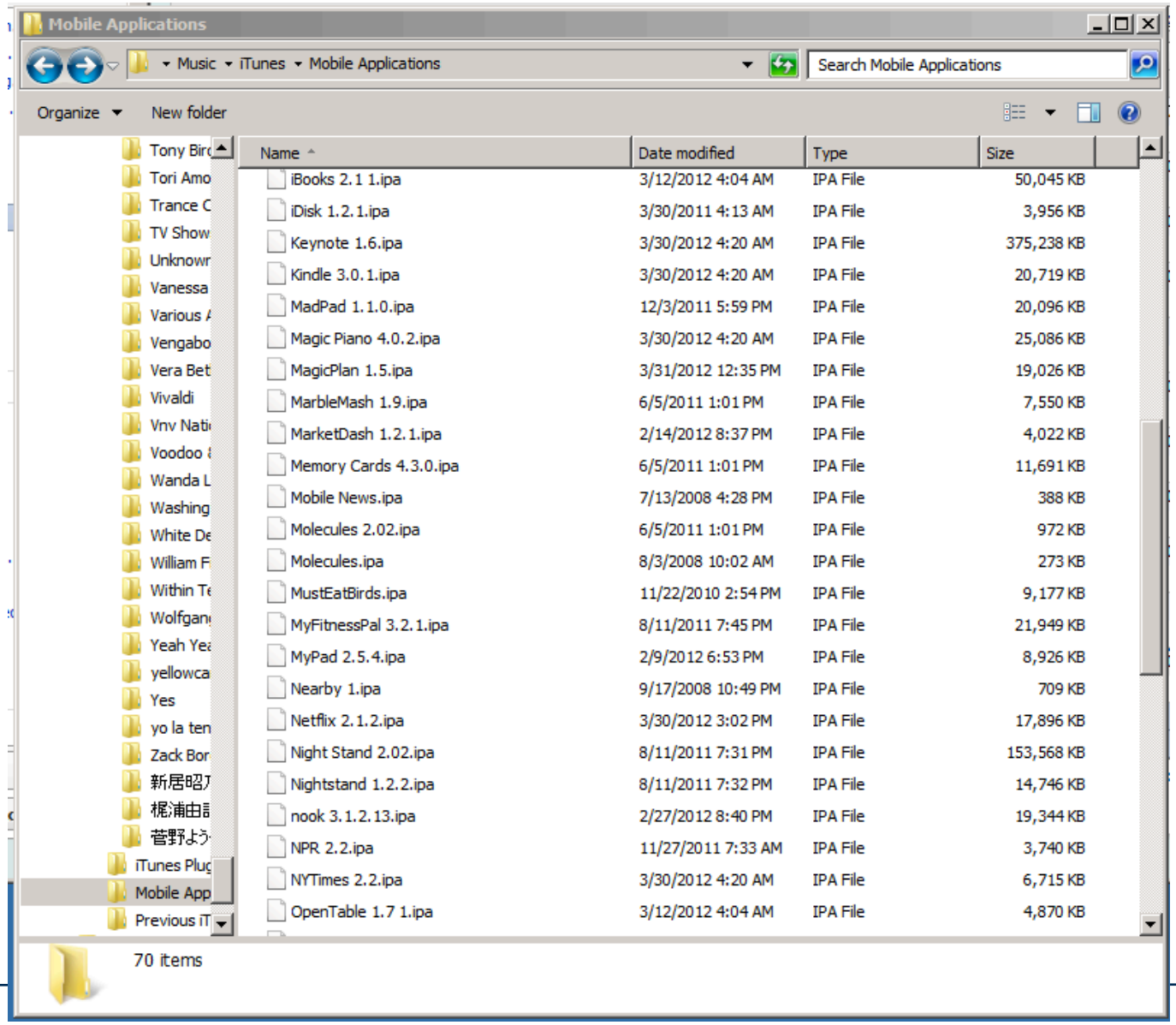
The challenge is identifying *likely voters.*

Hard drive contents can be predicted by sampling thousands of sectors:



The challenge is *identifying the sector* content that is sampled.

We think of computers as devices with *files*.

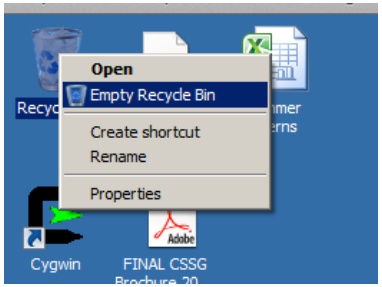


Data on computers is stored in fixed-sized sectors.

Data in a sector can be resident:

user files
email messages
[temporary files]

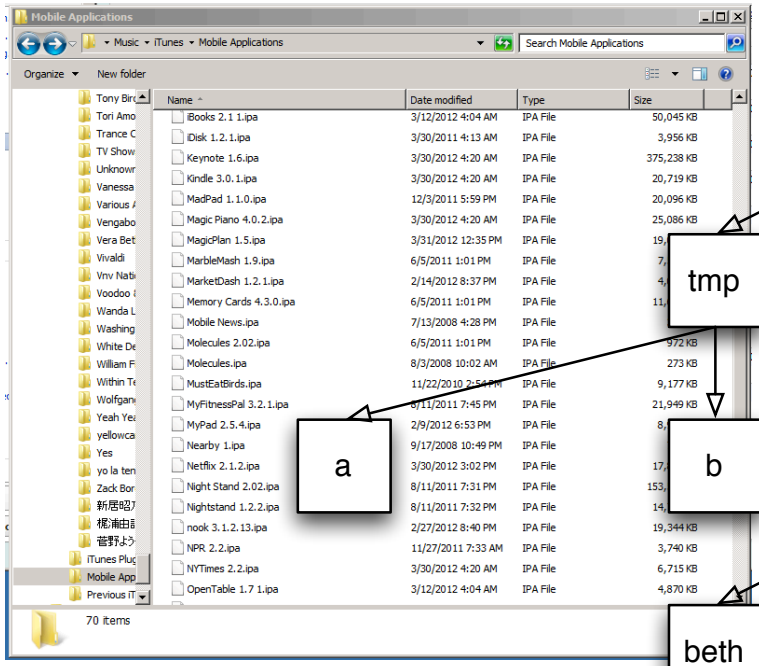
Files can be “deleted” but the data remains:



Sectors can be wiped clean:

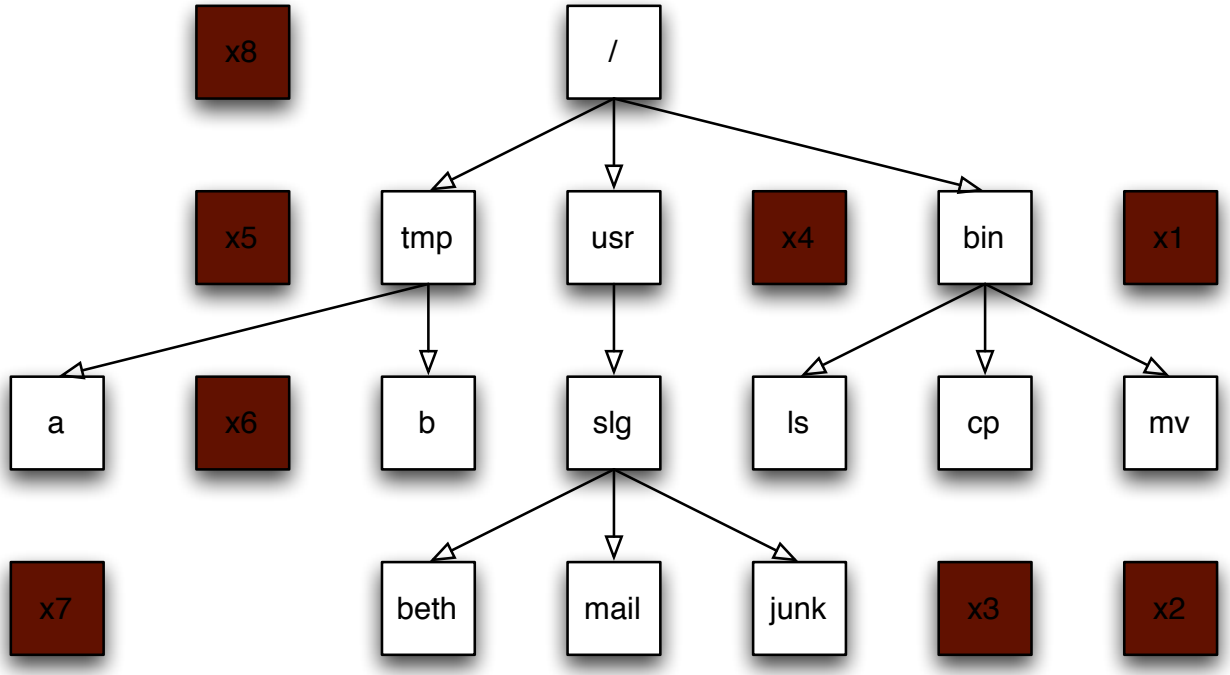
blank sectors

Allocated data are the data you see from the root directory. e.g. “visible” files.



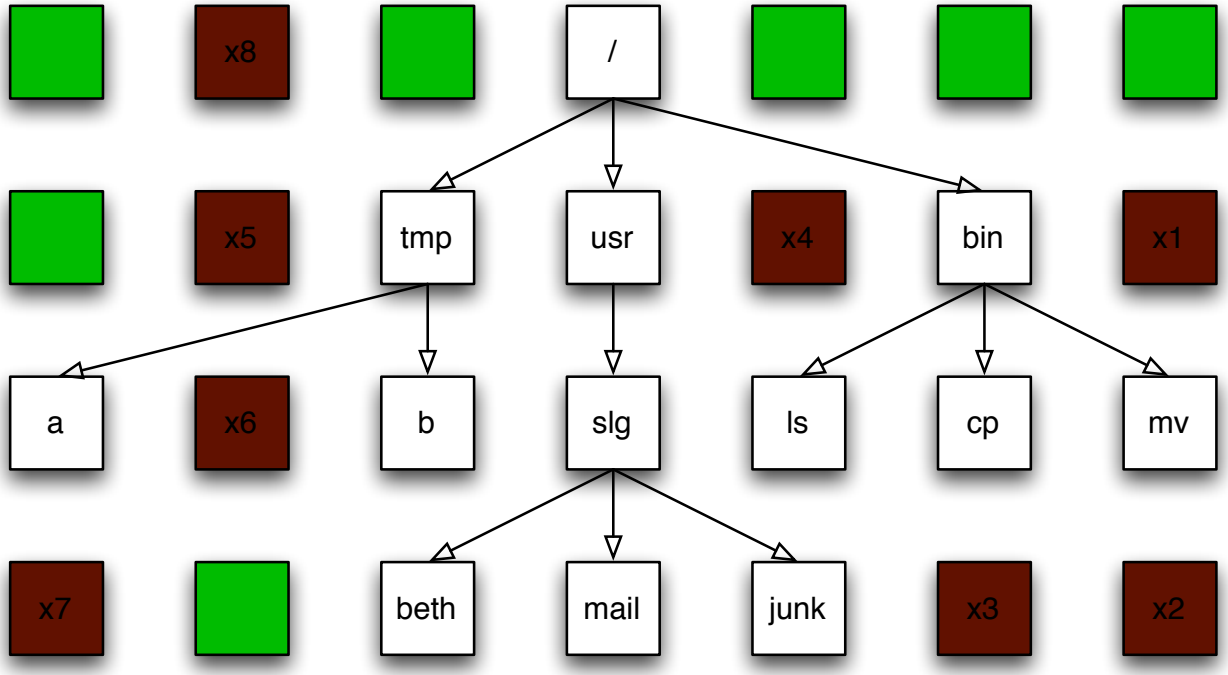
Resident Data

“Deleted data” are on the disk, but can only be recovered with forensic tools.



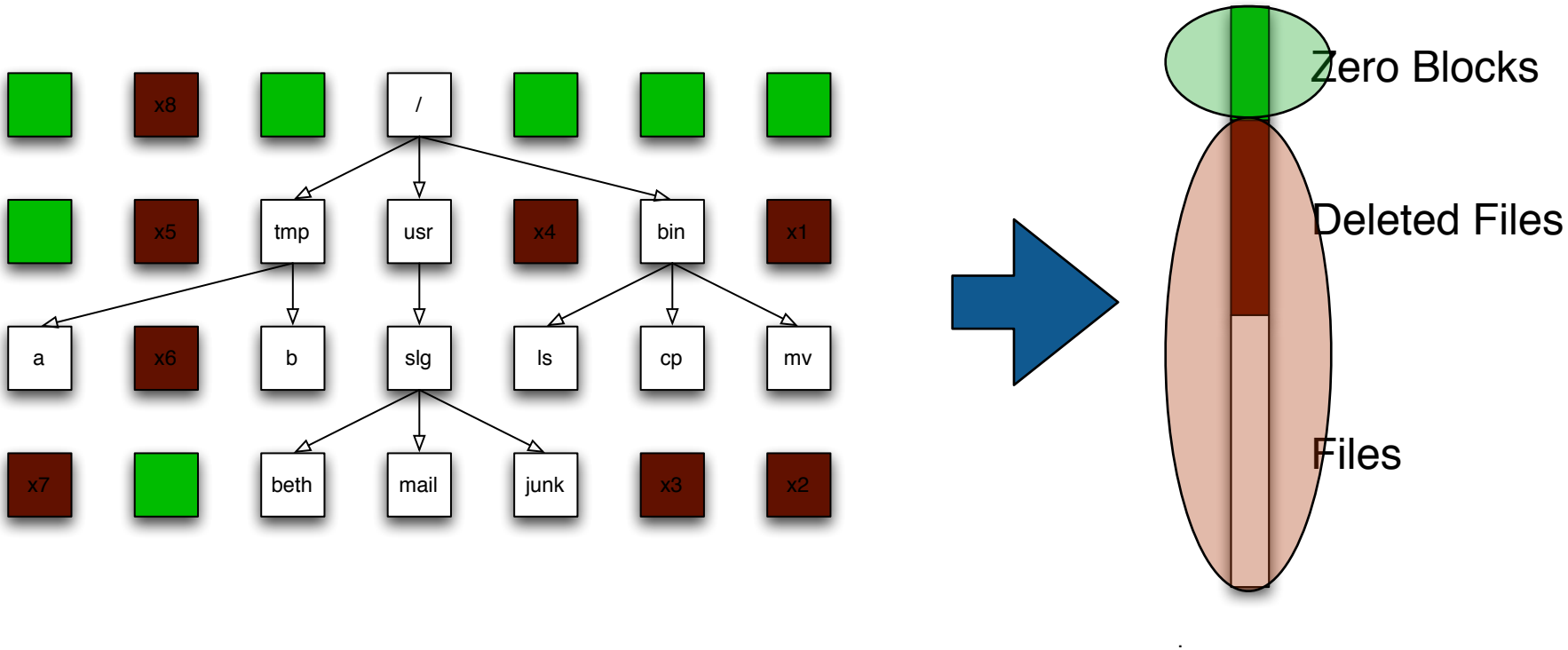
Deleted Data

Some sectors are blank. They have “no data.”



No Data

Sampling can't distinguish *allocated* from *deleted*.

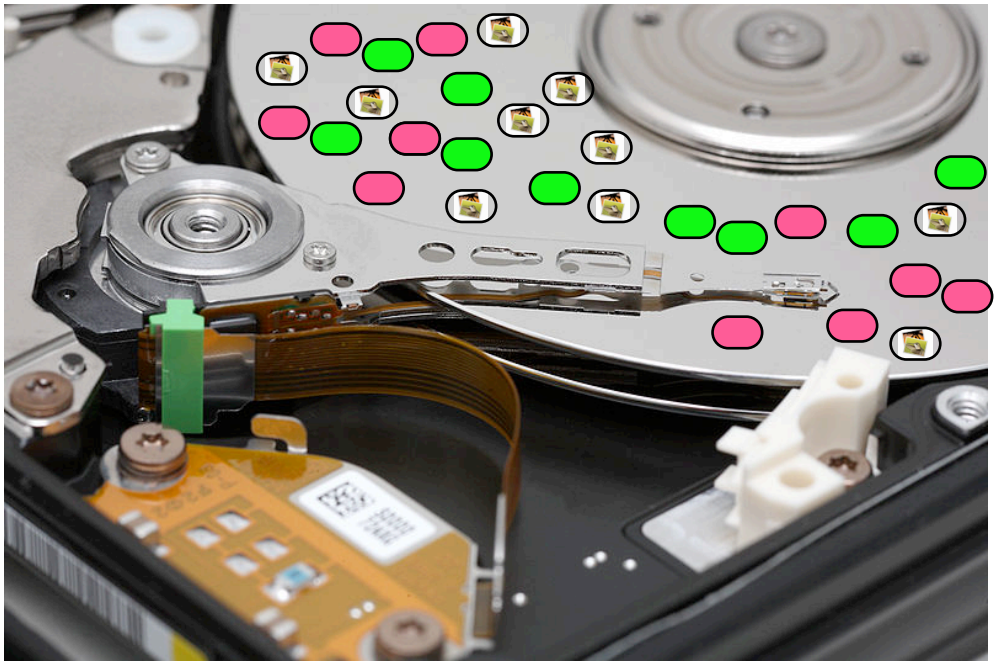


Sampling can tell us about the content of the data

Sampling can tell us the proportion of...

—*blank sectors; video; HTML files; other data types...*

—*data with distinct signatures...*



...provided we can identify it

Challenge for sampling: interpreting each sector

—*Easy:*

```
0000000: ffd8 ffe0 0010 4a46 4946 0001 0201 0048  ......JFIF.....H
0000010: 0048 0000 ffe1 1d17 4578 6966 0000 4d4d  .H......Exif..MM
0000020: 002a 0000 0008 0007 0112 0003 0000 0001  .*..............
0000030: 0001 0000 011a 0005 0000 0001 0000 0062  ...............b
0000040: 011b 0005 0000 0001 0000 006a 0128 0003  ...........j.(..
0000050: 0000 0001 0002 0000 0131 0002 0000 001b  .........1......
0000060: 0000 0072 0132 0002 0000 0014 0000 008d  ...r.2..........
0000070: 8769 0004 0000 0001 0000 00a4 0000 00d0  .i..............
0000080: 0000 0048 0000 0001 0000 0048 0000 0001  ...H.......H....
0000090: 4164 6f62 6520 5068 6f74 6f73 686f 7020  Adobe Photoshop
00000a0: 4353 2057 696e 646f 7773 0032 3030 353a  CS Windows.2005:
00000b0: 3035 3a30 3920 3136 3a30 313a 3432 0000  05:09 16:01:42..
00000c0: 0000 0003 a001 0003 0000 0001 0001 0000  ................
00000d0: a002 0004 0000 0001 0000 00c8 a003 0004  ................
00000e0: 0000 0001 0000 0084 0000 0000 0000 0006  ................
00000f0: 0103 0003 0000 0001 0006 0000 011a 0005  ................
```

—*Hard:*

```
000a000: 0011 fa71 57f4 6f5f ddff 00bd 15fb 5dfd  ...qW.o_......].
000a010: a996 0fc9 dff1 ff00 b149 e154 97f4 efd5  .........I.T....
000a020: e3f5 7f47 71df 8ffb d5d7 da9e d87f c12f  ...Gq........../
000a030: f8ff 00d8 b1f4 b1f8 ff00 c57e ab7a ff00  ..........~.z..
```

We use two approaches for identifying data type.

1 - SVMs with multiple feature types

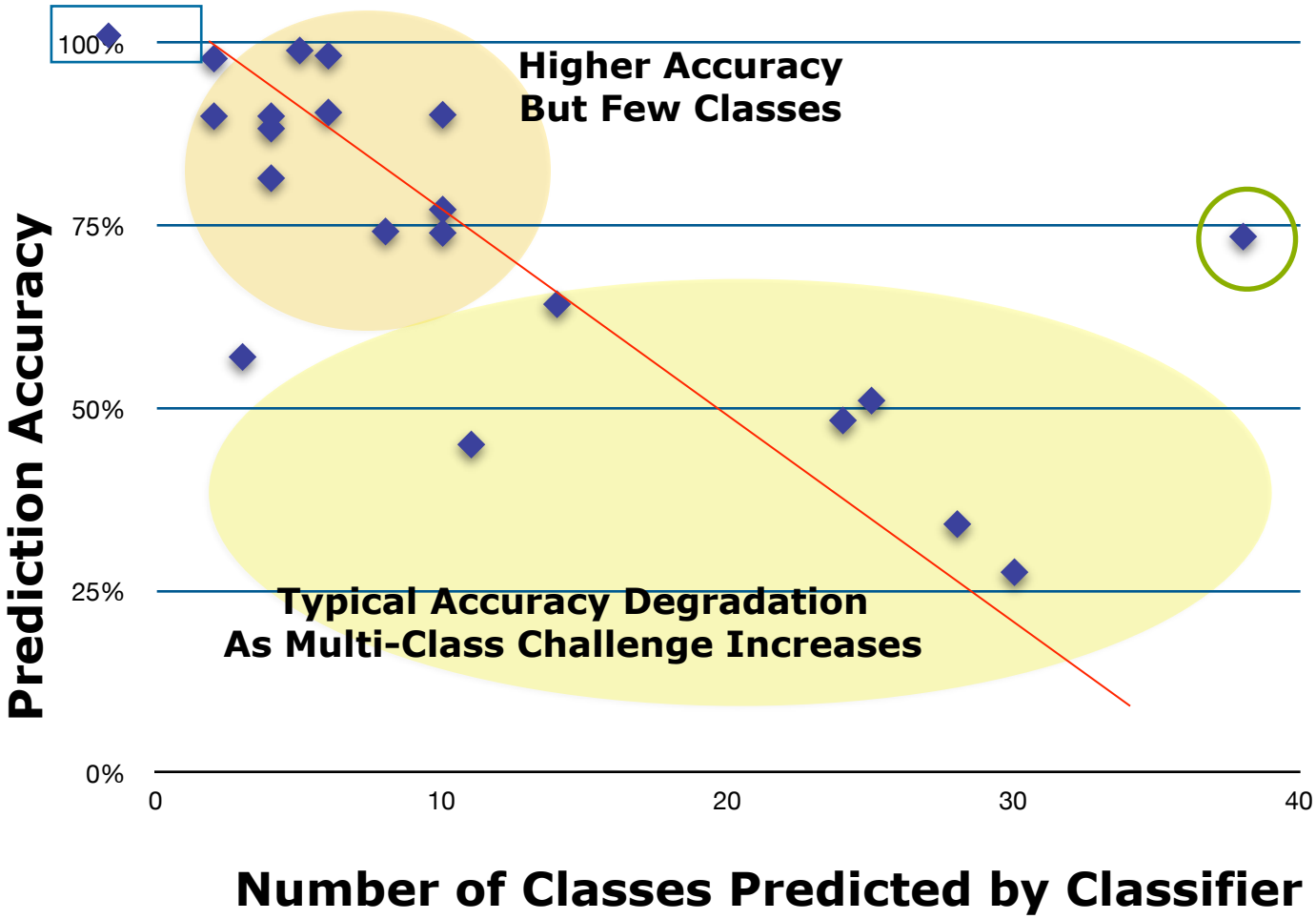
- unigrams
- bigrams (selected)
- Other n-grams & complexity measures
- compressibility
- hand-tuned classifiers

Beebe, N.L.; Maddox, L.A.; Lishu Liu; Minghe Sun, "Sceadan: Using Concatenated N-Gram Vectors for Improved File and Data Type Classification,"Information Forensics and Security, IEEE Transactions on , vol.8, no.9, pp.1519,1530, Sept. 2013

2 - Known content

- Database of “sector hashes.”

Sceadan provides the “type” of fragments.



Sceadan v1.0
73.5% Accuracy*
40 Classes

NOTE: 9 lowest performing types are significantly under-researched classes (e.g. Office2010, FS data)

***Additional model training has improved classifier accuracy from 71.5% to 73.5%**

Improved performance comes from feature set. Training is slow, but only needs to be done once.

Trigrams proved most accurate (70.19%)

- Much slower prediction time than competing alternatives

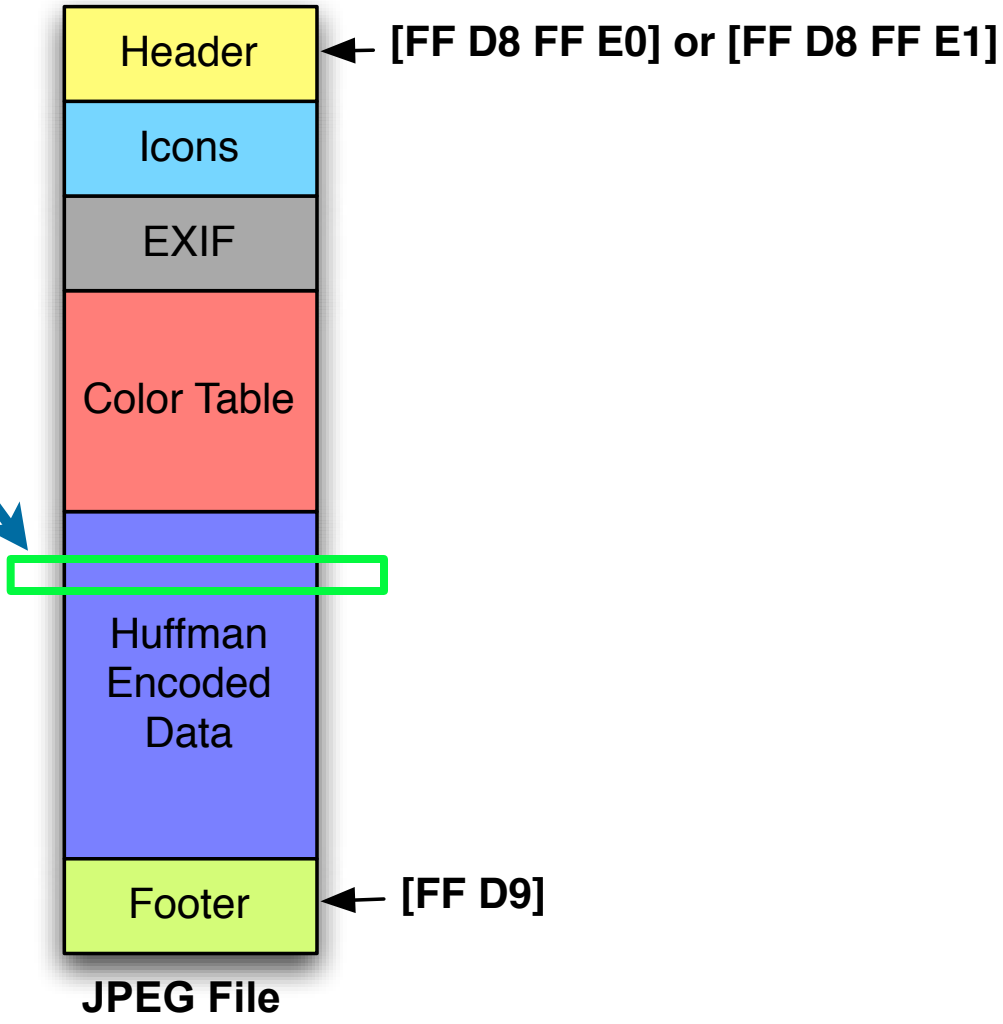
"FS5" (feature set 5) nearly as accurate (69.83%)

- Unigrams+Bigrams+Other
  - —*Other features: entropy, Kolmogrov complexity, mean byte value, Hamming weight, avg. contiguity between bytes, longest byte streak*

	S2			S3			c
Features	Train.Time	Pred.Time	accuracy	Train.Time	Pred.Time	accuracy	
unigrams	19m 9.518s	4.208s	55.99%	29m 46.439s	4.162s	48.20%	256
bigrams	5h 22m 21.391s	31.286s	68.12%	4h 34m 39.545s	32.649s	68.26%	1024
trigrams	174h 46m 5.795s	7m 47.676s	62.76%	211h 8m 47.311s	7m 23.068s	70.19%	1024
uni+bi	7h 39m 46.240s	36.019s	68.68%	3h 54m 36.043s	37.834s	67.06%	256
FS5	7h 51m 35.550s	35.111s	69.83%	7h 27m 34.618s	36.697s	68.92%	256

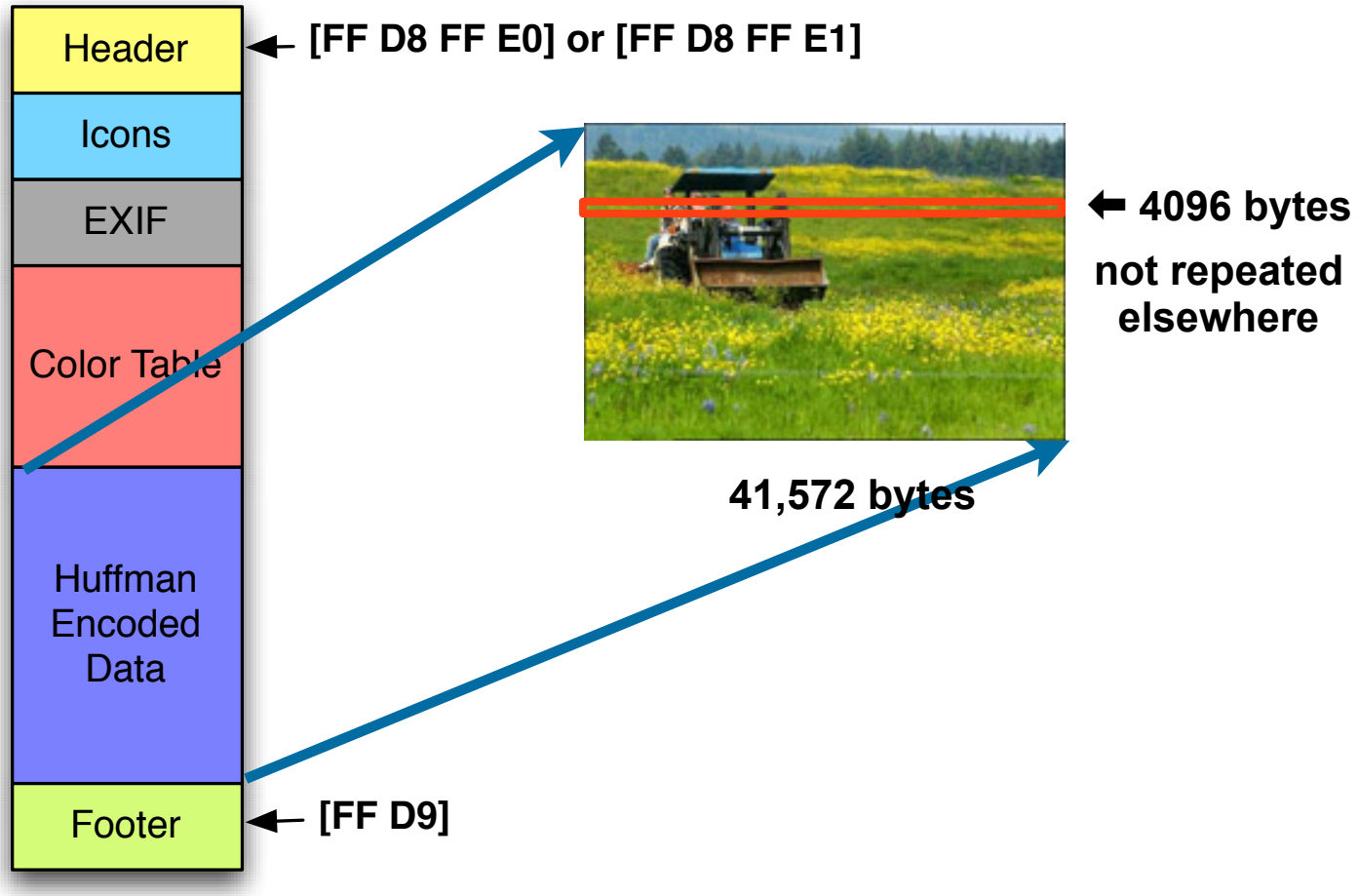
Some kinds of files have distinct contents.

Can you identify a JPEG file from reading 4 sectors in the middle?



We can identify “distinct” sectors.

In a compressed or encrypted file, each sector is different.



Initial anomaly detection results are promising.

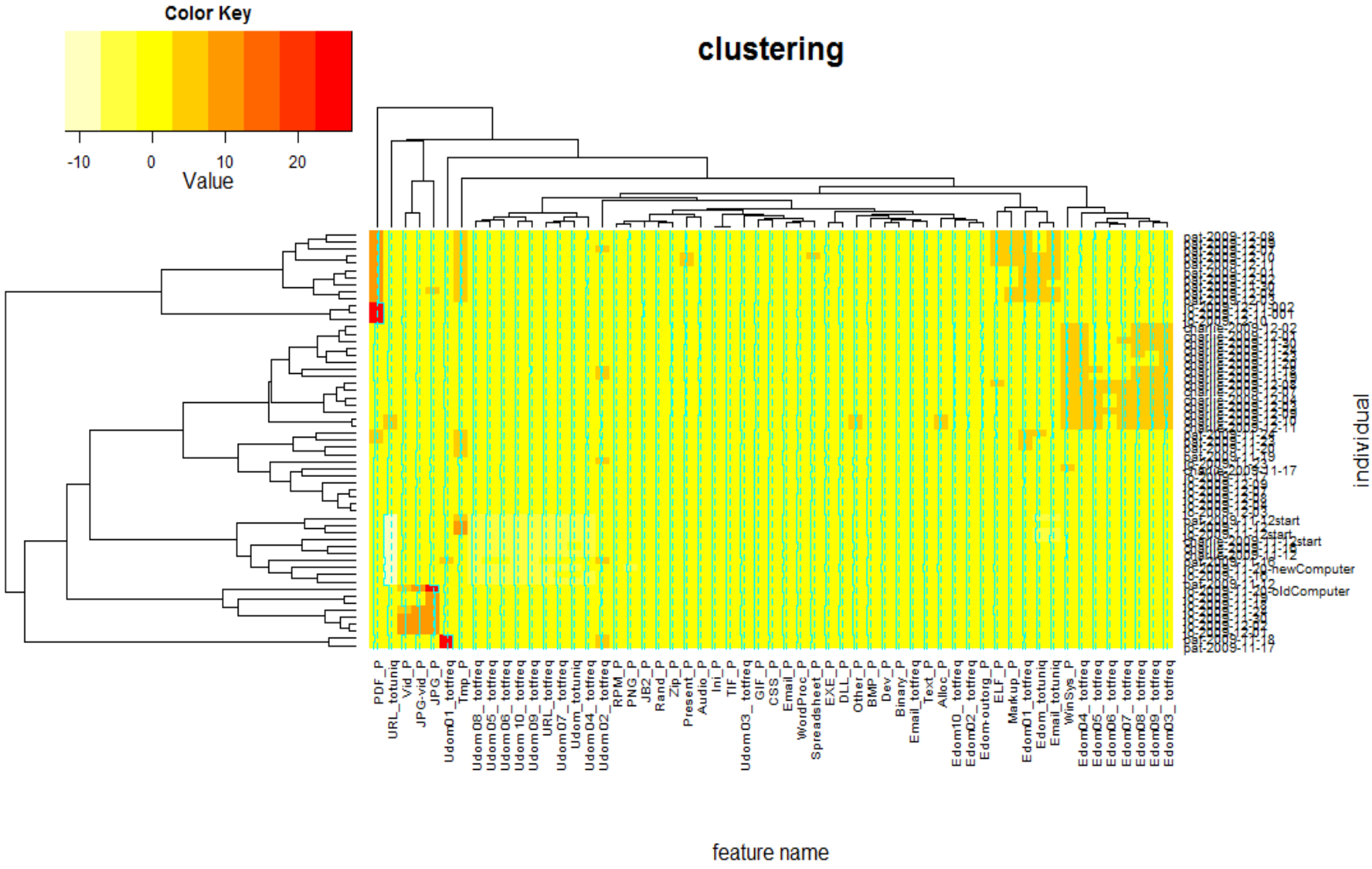
Successfully detecting univariate outliers

- Data by type most effective thus far
  - —*File types (e.g., jpg, exe)*
  - —*Data types (e.g., PII, CCN)*
- Median absolute deviation (MAD) based outlier detection with conditional scaling procedures

Multivariate and time-series based outlier detection — on-going

- Cluster based, SOM based, etc.

This heatmap of anomalies let an analyst easily identify clusters and outliers.



Current status — We're making progress!

bulk_extractor updated v1.4 just released

- Added features & GRR integration preparation

Sceadan data type classifier updated v1.2 released

Extraction, transformation, loading of datesets

- M57 Patents (digitalcorpora.org) case

Progress on anomaly detection algorithm

- Real Data Corpus extraction, translation and loading near complete
- Theoretical development
- Empirical data descriptive analyses (test assumptions)
- Univariate anomaly detection performing well on synthetic data set

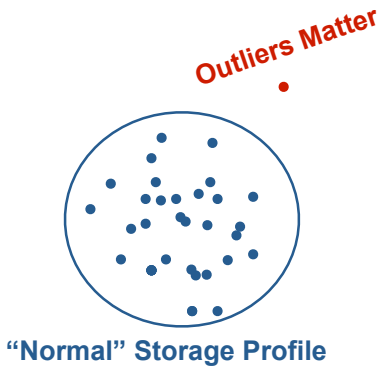
We are in year 1 of a 3-year effort.

	NPS Lead	UTSA Lead
Year 1	bulk_extractor upgrades	Outlier detection algorithm Synthetic data experimentation Real Data Corpus experimentation
Year 2	Integrate GRR Develop/test management console	Develop/test data outlier detection Develop/test visualization component
Year 3	Large-scale testing on partner net	Final dev. of outlier detection algorithm Final dev. of visualization agent

Many challenges remain.

"Anomalous" suggests "normal" exists

- Large, diverse, dislocated organizations
- High fluidity and variety in workforce
- Remote, mobile, multi-device access requirements
- Uninterruptible, critical computational operations



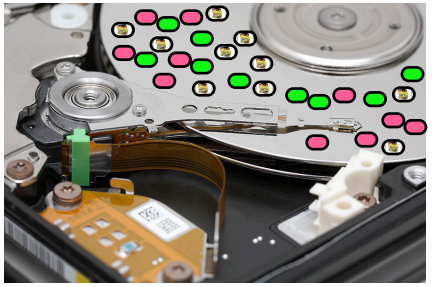
Clustering algorithm selection/development

- Accuracy and speed trade-off of extant algorithms
- Develop combinatorial algorithm to improve accuracy
- Need for automated parameter selection amidst noise
- Feature selection

Engineering of visualization component

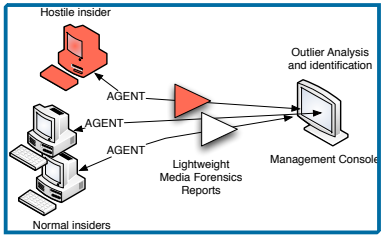
In conclusion, we are developing a system that uses "lightweight media forensics" to find hostile insiders.

We use random sampling to build a storage profile of media



We collect these profiles on a central server

We cluster & data mine to find outliers.



Contact:

- Simson L. Garfinkel simsong@acm.org
- Nicole Beebe *Nicole.Beebe@utsa.edu*