

Good morning thanks for coming to my talk. The slides will be posted to the NIST CFTT website in the near future. www.cftt.nist.gov

Disclaimer

Certain trade names and company products are mentioned in the text or identified. In no case does such identification imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the products are necessarily the best available for the purpose. No financial interest.

AAFS - Las Vegas, Nevada

Feb 25, 2016

2

I will try not to mention any specific products in my talk. If I do mention something I do not have any financial interest in any of these products.

The Problem With Characterizing the Reliability of Digital Forensics Tools

- Digital Forensic practitioners are confident that tools and methods are reliable
- Other forensic disciplines use error rates to describe chance of false positive, false negative or otherwise inaccurate results
- Confusion arises over the statistical use of the term *error* (a measure of uncertainty) and the day to day usage (a blunder or mistake)
- The court wants to know if results are reliable

AAFS - Las Vegas, Nevada

Feb 25, 2016

3

The court wants to know if the results presented are reliable. We know that our results are reliable. How can we communicate this to the court. Other disciplines can use error rates to describe the chance of false positives false negatives or otherwise inaccurate results but we do not always have that. The term error often causes a problem because the statistical meaning is a measure of uncertainty while the day-to-day usage is a blunder or mistake. This talk is based on the SWGDE document establishing confidence in digital forensic results by error mitigation analysis.

Guidelines, Not Rules

- Daubert - criteria to help assess reliability admissibility of scientific testimony
 - Tested
 - Peer review
 - Error rate
 - Standards & controls
 - General acceptance
- Daubert, Kuhmo Tire & GE v. Joiner.
- FRE 702

AAFS - Las Vegas, Nevada

Feb 25, 2016

4

Remember these are guidelines and not rules it's nice to be able to meet all of them but you don't have to.

Some Other Forensic Disciplines try to Match two Samples

- Fingerprint matching:
 - Suspect vs crime scene
 - Suspect vs data-base
- Same for DNA
- Tire tread
- Foot prints
- Tool marks & ballistics

AAFS - Las Vegas, Nevada

Feb 25, 2016

5

Other disciplines often focus on a single task such as matching one sample from the crime scene and a sample from a suspect.

Trying for a Match

- A technique declares a match or not
- The result and reality agree or not

And we get the usual 2x2 result table with type I and type II errors

Statistical analysis can give error rates

Matching is a natural for error rates

Testing a Hypothesis - Does entity X have attribute A?

- Statistical process, assumptions about randomness
- A Matrix of possibilities

Test Result	Reality	
	X has A	X does not have A
X has A	Accept	False Positive aka Type I Error
X does not have A	False Negative aka Type II Error	Reject

Error rate for each type of error is the probability of the error occurring.

AAFS - Las Vegas, Nevada

Feb 25, 2016

7

Matching is like a hypothesis test. Reliability can be measured with probability.

Digital Usually has more Questions

- Simplest question is: do two files match?
- Other questions:
 - Time line of events
 - Event reconstruction
 - Searching for strings
 - Document retrieval
 - Identifying file types
 - Recovering deleted files
 - Identifying deleted software

AAFS - Las Vegas, Nevada

Feb 25, 2016

8

A digital investigation is more than a single test.

Error Rate For Hashing Algorithm e.g., MD5, SHA1, Sha256, etc

- Two possible errors:
 - Two different files with different content & same hash
 - Chance of file collision
 - Error Rate is really small – practically zero
 - Two identical files with different hashes
 - can't happen
 - error rate is zero

Let me start with the easy example. Hashing algorithms have a built in chance of a false positive error that is unimaginably small. The algorithm is immune to false negative errors.

Comparing Randomly Selected Files

Chance of hash or checksum for matching any two files

Algorithm	Chance of Collision
CRC-16	1 in 32,768
CRC-32	1 in 2,147,483,648
MD5 (128 bits)	1 in 170141183460469231731687303715884105728
SHA-1	1 in 2^{159}
SHA-256	1 in 2^{255}

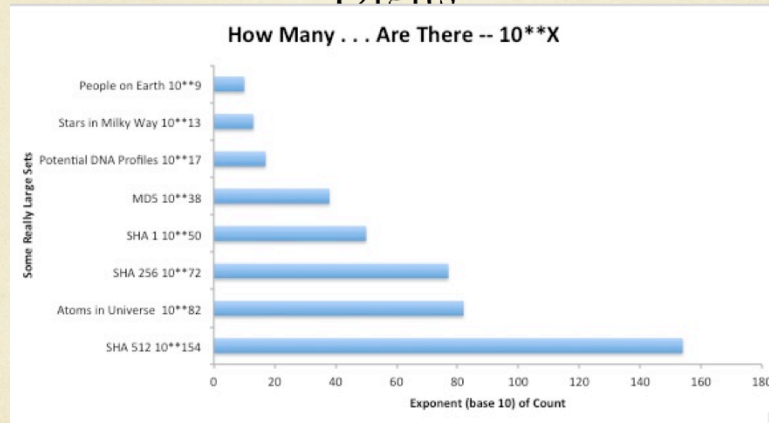
AAFS - Las Vegas, Nevada

Feb 25, 2016

10

The CRC checksums lack some desirable properties of the cryptographic hash algorithms like randomization of the output so that similar files (even one bit different) produce very different cryptographic hashes.

Some Big Numbers Graph of Number of Decimal Digits



AAFS - Las Vegas, Nevada

Feb 25, 2016

The probability of a hash collision is unimaginably small. MD5, considered “not good enough” by some, has a chance of hash collision better than one in the number of people that there would be if every star in the milky way galaxy had 10 planets with earth size populations ($10^{**}9 \times 10^{**}13 \times 10$ is only $10^{**}23$, this is far less than $10^{**}38$).

SHA512 is just overkill that’s been overkilled.

But an Implementation may have an error

- Not random in nature - rerun and get exactly the same result for the same input
- Systematic in nature - triggered by some conditions
- Example: MD5 hash program
 - Always correct running on Linux
 - If run in Windows, correct for binary files, fails for text files (Windows adds a line feed character at the end of each line)

12

AAFS - Las Vegas, Nevada

Feb 25, 2016

Here comes the rub, and it applies to any forensic process that uses computer software to calculate a result. A hypothesis test or a probability value depends on a random variable with a known probability distribution (usually Gaussian, aka Normal). The (random) error rate is a measure of uncertainty.

The software that makes the calculation can have a software error that is not random in nature. This is a systematic error, nothing random here. Same input yields same output.

BTW, I wrote this program on Linux and moved the software to windows. The software error quickly showed up in just a few test cases and was promptly fixed.

Not So Fast- More to the story

The court wants to know if testimony is reliable. What is the whole picture:

- Algorithm: Is it scientific/reliable/repeatable?
- Implementation: Does the software work?
- Application: Correct procedure followed?
- Interpretation: Did the examiner understand the result?

13

AAFS - Las Vegas, Nevada

Feb 25, 2016

The algorithm has an error rate, but the tool may have systematic software errors and there are other broad paths to perdition. A practitioner might not follow the best practice and wind up comingling data from two cases. Or a practitioner may think that a file was accessed at 00:00 (midnight), but in reality it was zero because the “access” field was never updated by that particular OS.

Sources of Error

- The theory of measurement error identifies two classes of errors: measurement (random process) & systematic (non-random)
- For forensic tools that implement some algorithm . . .
 1. An algorithm may have a theoretical (random process) error rate
 2. An implementation of an algorithm may have systematic (non-random) errors, i.e., software bugs
 3. The application of a procedure may have a blunder that affects the result
 4. A practitioner may misunderstand something
- The court wants to know that the final result is reliable.

AAFS - Las Vegas, Nevada

Feb 25, 2016

14

Here is a little clarification on the word error

Statistical vs systematic

Again, the court wants to know the result is reliable

Typical Errors in Forensic Tools

- Incompleteness – missed something
- Inaccuracy – something is wrong
 - Reported item does not exist
 - Reported item is altered, e.g., update time stamp
 - Association of unrelated items
 - Recognize corruption

15

AAFS - Las Vegas, Nevada

Feb 25, 2016

It helps to find errors if you can identify likely errors and then test for them.

These are the kinds of errors we have seen at CFTT

Error Mitigation Strategies

- Define likely errors & risks
- Test tools for likely errors
- Use written procedures
- Document observations, history of problems
- Oversight, Technical & Peer review
- Context Analysis of results – sensible answer

AAFS - Las Vegas, Nevada

16
Feb 25, 2016

With all the sources of error, what to do . . .

Test for the software errors that are likely or you have seen before

Follow good quality assurance procedures and best practices (like published by SWGDE)

What Do Digital Tools Do?

- Collection - disk imaging & write blocking
- Search - look for items that have properties of interest
- Reconstruction - Put things back together
- Time Line - When did events happen

17

Write Block Device Test Example

- Write blocker for either IDE (ATA) or SATA drives with host interfaces: SATA, USB, FW400 & FW800
- Need eight separate test runs: 2 drives x 4 interfaces (Can be tested in 30 minutes)
- Result:
 - All ATA commands blocked
 - All SCSI commands to FireWire blocked
 - “WRITE 16” NOT Blocked for USB (Only needed for drives larger than 2.1TB)

AAFS - Las Vegas, Nevada

Feb 25, 2016

18

Here is an example of what testing can reveal

There are about 5 write commands that a disk driver can choose from. A disk driver (software to access a storage device) usually has a preferred instruction for a given type of drive. In this case, on Windows XP, the write 10 command is preferred unless a disk address greater than 1.2TB is accessed. The “write 10” command has an address limit at that point and a command, like “write 16”, with a larger address range must be used.

Note that this particular write block device works just fine except for “write 16” over the USB interface. “Write 16” is blocked on the firewire interface. The problem arose when a chip maker implemented a significant change without informing the write block vendor.

File Recovery

- Different algorithms (different results)
- No one “right answer”
- Need to define error carefully
- Behaviors observed in recovered files:
 - Data from multiple files
 - Missing data (available but missed)
 - Overwritten data (overwriting data returned)

19

AAFS - Las Vegas, Nevada

Feb 25, 2016

File recovery is one of the more challenging tasks

Recovered files need to be checked for mixing data clusters from multiple files together

Graphic File Carving Behaviors

Success measured by ability to view returned file

- Beginning of file returned
- Only viewable in some file viewers
- Only one file viewable but additional graphics included in file
- File not viewable, only one sector missing
- Risk that recovered data already on storage device before used by current owner

20

AAFS - Las Vegas, Nevada

Feb 25, 2016

Viewing a file usually makes any mixing of data from multiple sources stand out and easy to identify

Summary & Observations

- Distinguish between intended algorithm and actual implementation
- Algorithm may have an error rate (statistical in nature)
- Implementations have systematic errors
- Most digital forensic tool functions are simple collection, extraction or searching operations with a zero error rate for the algorithm.
- Tools tend to have minor problems, usually omitting data, sometimes duplicating existing data.
- An implementation's systematic errors can be revealed by tool testing.
- To satisfy the intent of Daubert, tools should have the types of failures and triggering conditions characterized.
- Error mitigation analysis involves recognizing potential sources of error
- Taking steps to mitigate any errors
- Employing quality assurance and continuous human oversight & improvement

AAFS - Las Vegas, Nevada

Feb 25, 2016

The key message from the SWGDE document is to look at Error holistically – examine what kinds of errors can occur, which ones are likely. Then systematically take steps to address and reduce error and to describe where potential errors (especially the likely ones) remain.

References

- See: SWGDE *Establishing Confidence in Digital Forensic Results by Error Mitigation Analysis* at www.swgde.org
- CFTT: www.cftt.nist.gov

22

Contact Information

Jim Lyle
jlyle@nist.gov

Sue Ballou
Susan.ballou@nist.gov

Barbara Guttman, Software and Systems Division
bguttman@nist.gov

AAFS - Las Vegas, Nevada

Feb 25, 2016

23