LECTURE 1

WHAT IS DIGITAL FORENSICS?

Application of science to identification, collection, examination, and analysis of data while preserving the integrity of the information and maintaining a strict chain of custody for the data.

- If the data is damaged, stolen or compromised the organization needs the capacity to reconstruct what occurred.

DIFFERENCE BETWEEN CYBERSECURITY AND DIGITAL FORENSICS

- **CYBERSECURITY:** Monitor and prevent unauthorized access, misuse or modification of digital data, a device or a network.
- **DIGITAL FORENSICS:** Monitoring and analysis of the digital evidences with the goal of reconstructing its story and understand what happened.

PROBLEMS FOR DF

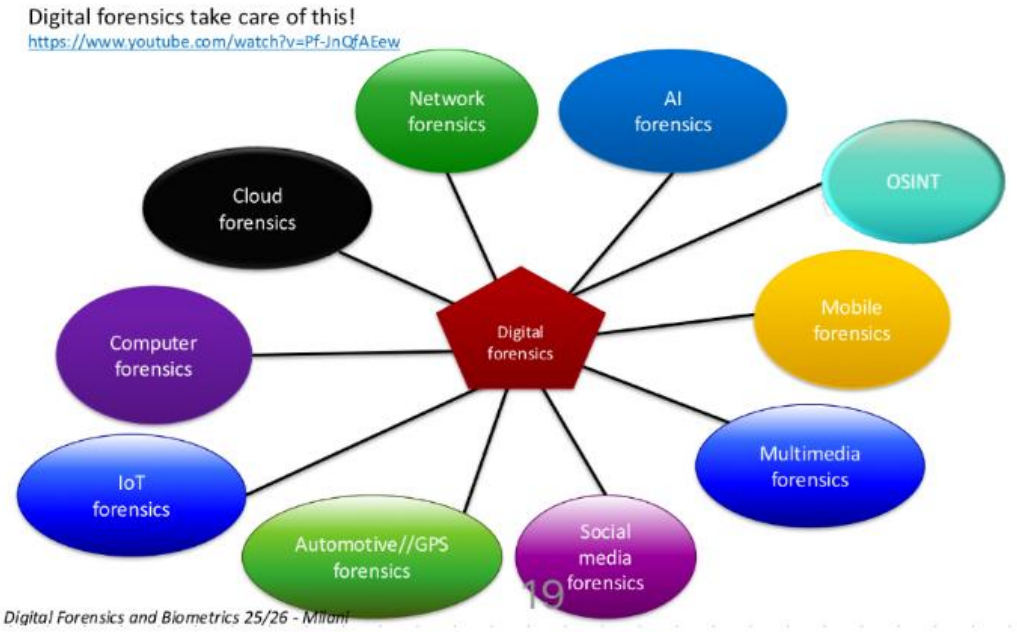
- Who perpetrated the act? Is there attribution?
- What did they gain from the attack?
- What did we lose?
- Where did it happen? On a server or a host?
- When did the exploit execute? Over what time span?
- How did they execute the exploit?

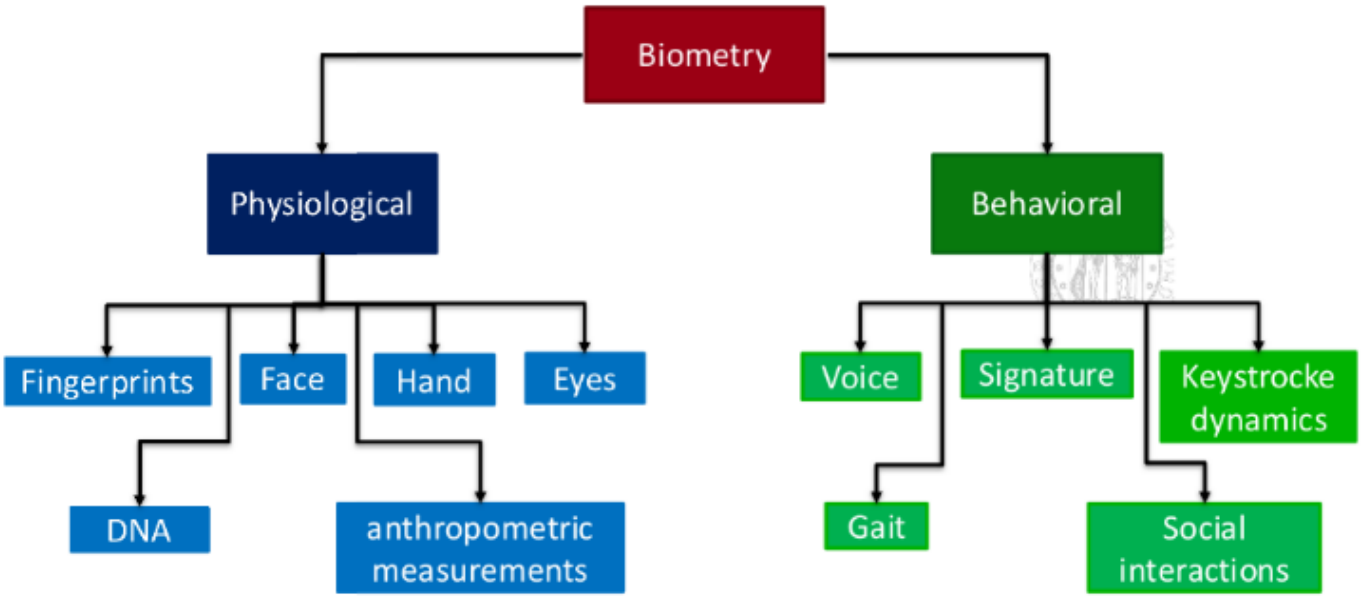
Important: **Digital ID** and **Biological ID.**

AUTHENTICATION MECHANISMS

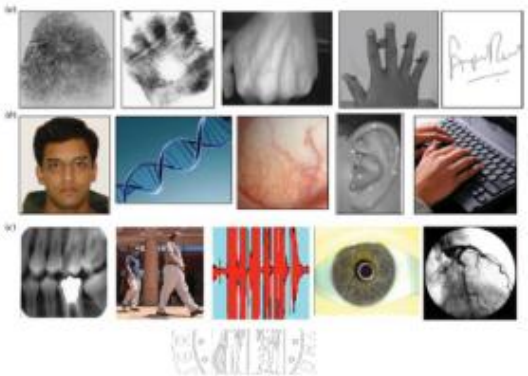
- **Something u Know** (password or PIN)
- **Something u have** (cards or tokens)
- **Something u are** (biometrics)

Biometrics: Measurement of physical characteristics of personal traits.





Who or what is the source (person) related to this evidence?



Possible outcome:

- Individualization
- Inconclusive
- Exclusion

Associate an evidence (measure) to an individual but

1. Forensic analysis is applied after the event has occurred; biometrics can be applied before the event (access control).
2. In a forensic investigation you can not control in advance the type of measurement
3. Forensic evidence collectionis mostly manual
4. Biometric recognition probably needs to be done in real time
5. In forensics, false negative is to be minimize: exclude the perpetrator. In biometrics, it depends: in surveillance, false non-matches need to be minimize w.r.t. false matches; in access control, avoid false matches
6. Inconclusive can be compensated in biometrics by additional acquisition
7. In forensics, measures have a lower quality
8. The outcome of a forensic investigation has to be declared verbally to a jury or a court; in biometrics, outcomeis parameterized numerically

LECTURE 2

ACQUIRING DIGITAL EVIDENCE

PEOPLE AROUND DE

- **DE (Digital Evidence)**
- **DEFR (Digital Evidence First Responder):**
  - He is in charge and coordinates.
  - He can interact with IRS to get more information
  - **Duties:** Isolate the device, seize it, or eventually acquire some files. Coordinate the steps.
  - **Definition:** An appointed specialist intervenes on the scene to acquire the digital evidence: he can isolate the device, seize it or eventually acquire some files.
  - Authorized operator that accesses information devices and systems for the first time
  - Must be experienced
  - May use collaborators
  - **What does he have to do?**
    1. Check and seize the area containing digital devices.
    2. Identify the person in charge of that area.
    3. Keep people away from devices and power supply.
    4. Profile all persons authorized to access the area.
    5. **Crucial:** Avoid changing the state of devices.
    6. Document the scene, components, cables (pictures, video, drawings, schemes).
    7. Identify notes, diaries, post-its, manuals (look for passwords and PINs) – these are *not* digital evidence but are critical.
- **Steps:**
  - Evidence is brought to the lab for analysis
  - Forensic specialist comes into play
  - Conclusions can be drawn.
- **DES (Digital Evidence Specialist):**
  - Operator that is expert in digital evidences.
  - Different from DEFR. The different is that the DEFR is authorized, prepared and qualified to operate first on the crime scene to acquire and store digital evidences instead the DES can operate as a DEFR but has the competences to deal with a wide range of different technical issues (e.g., OS, network acquisitions, etc.)
  - He helps the DEFR.
  - He can operate as a DEFR but has higher technical competences to deal with a range of different technical issues.
- **FLM (Forensic Lab Manager):** He guides the lab (operates inside the lab).
- **IRS (Incident Response Specialist):** Operates first "after the fact" (it can't be someone in charge of the investigation).

DIGITAL EVIDENCE ACQUISITION PHASES

The 4 Phases:

1. **Identification:** Search, recognition, and documentation of possible evidence.
2. **Seizure:** Acquisition of the physical device containing possible DE.
3. **Acquisition:** Creation of a copy of the data.
4. **Preservation:** Storage and conservation of the integrity and original condition of the potential DE.

Storage Space:

- Must be a safe environment where DE are kept.
- Storage devices must not be exposed to magnetic fields, dust, vibrations, or other agents (e.g., humidity, temperature) which could compromise the potential DE.

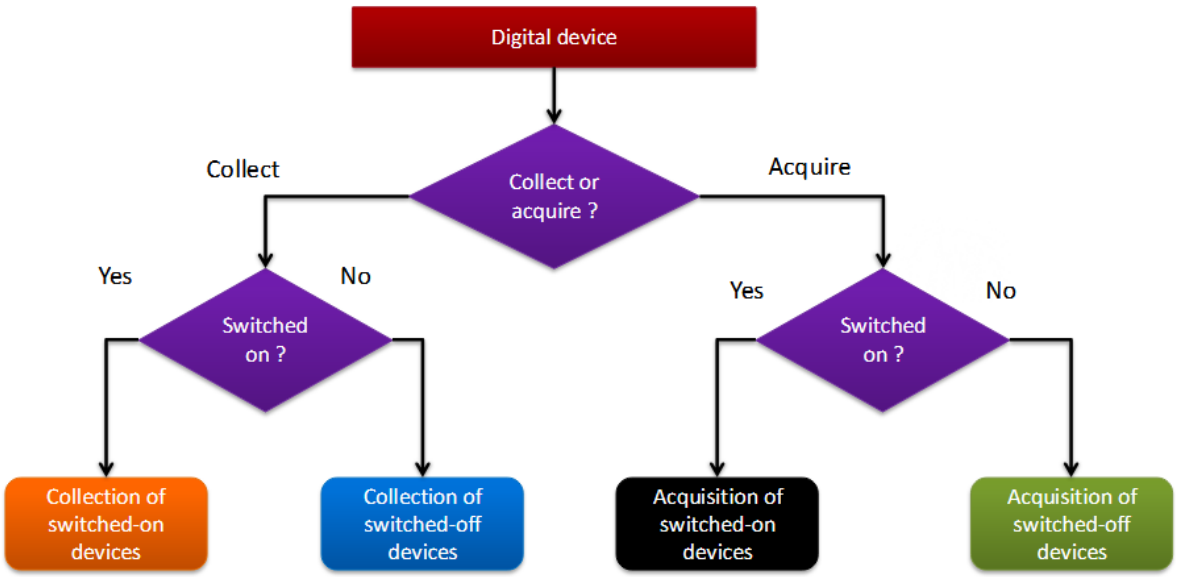
IDENTIFICATION & Volatility:

- A digital device has a physical and logical consistency.
- Some devices can present **Volatile Data**: Data stored in temporary memory (like RAM) that is lost when the system loses power or is restarted.
- **Network Note:** If a device has a network interface, we must identify which terminals exchanged data/comunications with it (regardless of the fact that it is connected or not at the intervention time).

ORDER OF VOLATILITY (ODV):

1. CPU Register, Cache, and Real-time State.
2. RAM (Main Memory).
3. Temporary Memory in Network Devices.
4. Memory in Peripheral Devices.
5. Disk-based Volatile Data.
6. Remote and Cloud-based Volatile Data.

IDENTIFICATION: SCHEME



Difference between Collection and Acquisition:

- **Collecting:** Gathering and preserving potential devices.
- **Acquiring:** Creating a forensically sound copy of that evidence for analysis.

IDENTIFICATION: DEFR'S OPERATIONS

- Document type, brand, S/N for each storage device.
- Identify all computers and peripherals together with their state.
- Collect power cables for each device using batteries.
- Use a wireless signal detector to identify eventual systems that are present but not immediately visible.
- Take into consideration possible non-digital evidence.
- **Pay attention to:** Volatility, Cryptography (on devices or partitions), Critical elements, Legal requirements, Resources (availability of storage/time/personnel).

COLLECTION OR SEIZURE

Seizure/Collection vs. Acquisition:

- **Seizure/Collection:** Devices are seized and brought to the lab. Can be physically removed from the location.
    - Devices can be switched on or off
        - DEFR must use the best practices based on costs and times (to be documented)
        - Accessories must be gathered as well
- **Acquisition:** Logical extraction of information from the device (e.g., extract files).

COLLECTION: CRITICAL SITUATIONS:

There are devices that can not be switched off (Special attentions are required)
We can perform:

- **Live Acquisition:** Data are acquired on the spot. Needed when the device cannot be seized, or you fear losing information.
- **Partial Acquisition:** Only parts of the data are acquired. Used when:
    - System contains too much data (e.g., DB servers).
    - System cannot be switched off.
    - Only part of the data is relevant.
    - Only part of the data can be acquired (legal issues).

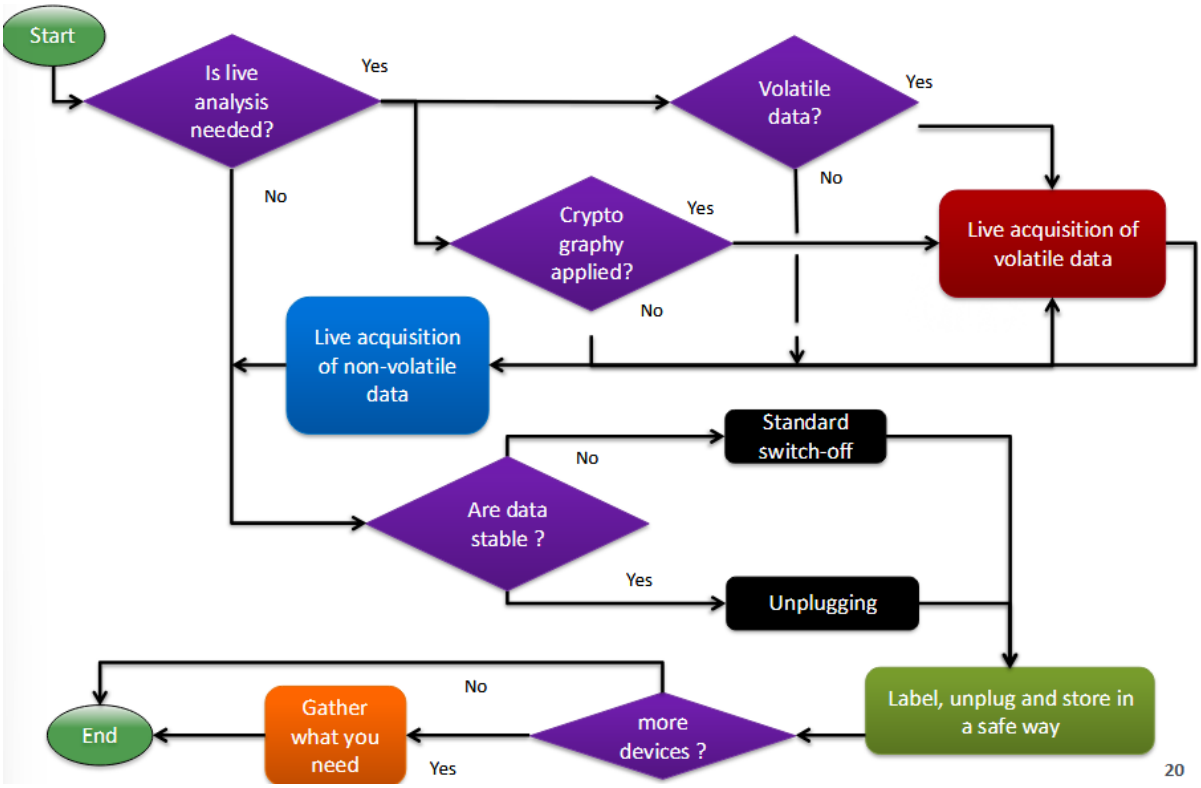
Decision Making Flowchart

Live Analysis & Switch-off Decisions:

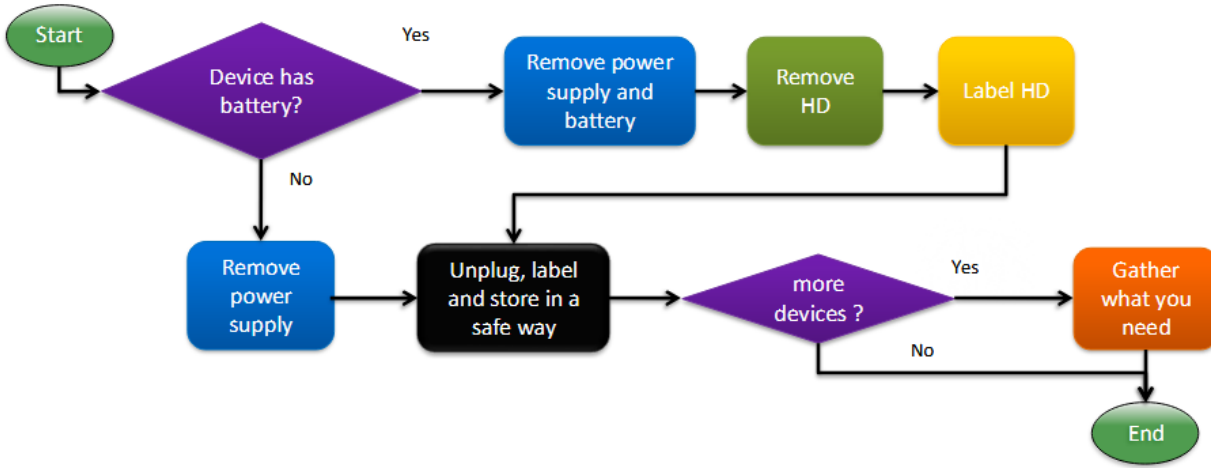
1. **Is Live Analysis needed?**
    - *Yes* -> **Is Crypto applied?**
        - *Yes* -> Live acquisition of volatile data.
        - *No* -> Live acquisition of non-volatile data.

- *No* -> **Are data stable?** (You must ask yourself this).
  - *No* -> Standard switch-off.
  - *Yes* -> Unplugging.

COLLECTION: SWITCHED-OFF DEVICES



COLLECTION: SWITCHED-OFF DEVICES



Procedures for Specific Devices

1. **Verify:** That all devices are really switched off.
2. **Remove:** Memory storage from device (if possible/applicable).
3. **Label:** Memory device and power cables.
4. **Document:** Details (type, vendor, S/N, part number, size).
5. **Acquire:** Compute hash value for integrity verification.

- *Note:* You can isolate a mobile device (phone) in case you can't remove the battery. -> the problem is that the device can run out of power.

ACQUISITION

Goal: Create a forensic copy and document the method, tools, and activity.

- **What to acquire:** Support, Partition, Files.
- **Rule:**
  - Avoid altering the material (don't change any bit).
  - If not possible, document modifications and justify.
  - Get data in a complete way minimizing alterations
  - Printing and copying can be done but most of the meta-data are lost
- **Imaging:** Creates a bit-stream copy of a device in a non-invasive, non-altering way. All parts of the disk are copied (including metadata).

Types of Copies:

- **Reliable Forensic Copy (all the meta data must be copy as well):**
  - Clone of a disk
  - Bit-wise copy image
  - Compressed bit-wise copy image
- **Cloning:** Creates a functional copy (disk-to-disk). It creates a "hard copy".
- **NOT Forensic:** Cut & Paste or Drag & Drop are *not* forensic copies because they delete/alter metadata (miss erased files, slack space, free space). (temporal data are lost)

NB. Disk cloning creates a functional one-to-one copy of a hard drive, while disk imaging creates an archive of a hard drive that can be used to make a one-to-one copy

Tools & Formats:

- **Linux:** dd command (e.g., dd if=/dev/sdc of=image.dd).
- **dcfldd:** is an advanced version of **dd**. Includes hashing algorithms and splitting input image (e.g., dcfldd if=/dev/sdc hash = md5,sha256 hashwindow = 10G md5log = md5.txt \ sha256log = sha256.txt split = 10G splitformat = aa of = imagel.dd).
- **Proprietary Formats:**
  - EnCase, ILook, non-compressed (IRBF), and encrypted (IEIF),
  - encapsulated metadata,
  - need appropriate program.
- **Independent File Formats:** (AFF, AFD, AFM)
  - can be used across different tools
  - collect disk and imaging process metadata

EXAMPLE OF LIVE ACQUISITION OF A POST

Acquire images from social media/web sites

Download with picture documentation (Es. screenshot or printout)

Problems:

- Malware
- Human errors
- Alterations

OS live (open source), Video registration, Visit some web sites: register network traffic, Use different softwares for hash computation
Profiling (SP not collaborating, shared PC)
Alternative: LegalEye
Problems: integrity of the environment and network
Given a file, **how can I ensure it has not been altered?**: Compute a sort of "signature" that changes whenever the file is altered.

HASH FUNCTIONS

- Images are widely accepted in courts.
- Create more copies: 1 master and more copies to be given to all the legal actors involved
- Permits to give the resouce back to the owner

Integrity must be verified after acquisition and generation of copies. Create a digest using a hash function:

- $H: file \rightarrow 0,1^{n}$

  $f \rightarrow\ d(f)$

Digest: string of symbols/bits with predefined length generated by a hash function operated on the bits of a file/image.
One-way: Hash function is non invertible function that maps variable length strings of symbols/bits into fixed length strings of bits

- H is not mutual
- From digests it is not possible to reconstruct files
- It is very unlikely that two different files has the same digests (although probability is not null: collisions)
- It can be used to verify errors

HASH FUNCTIONS FOR CRYPTOGRAPHY
Properties:

1. **Resilience to Pre-image:** given a digest ***d***, creating a variable-length string f s.t. ***d(f)=d*** must be computationally-unfeasible.
2. **Resilience to Second Pre-image:** given a digest $d_1 = d(f_1)$ computed on a variable-length string f1, it must be computationally-unfeasible to find f2 s.t. $d_2 = d(f_2) = d_1$.
3. **Resilience to Collision:** searching for a couple of strings f1 and f2 such that $d(f_2) = d(f_1)$ is computationally-unfeasible.

MD5 (Message Digest algorithm 5):

Message Digest algorithm 5 , RFC1321 Ronald Rivest, 1991

variable

$$MD5: \{0,1\}^{l} \rightarrow \{0,1\}^{128}$$

$$\mathbf{s} \mapsto \mathbf{d(s)}$$

Maps variable-length strings into strings with 128 bits

MD5 checksum or MD5 hash

- Fast
- Highly-improbable to find 2 different strings with the same MD5 hash
- Many resources online allows to decipher common words (www.md5online.org)
- Widely diffused

is possible to invert this if u have short lenght

Example: ➡ = decrypt!

"Digital Forensics" MD5: **093005a1f5950fba6c6a8b628c8e3a22** ➡ **"Digital Forensics"**

"This is ciphered text for Digital Forensics course!" MD5: **8bdf55f7b5ec03f5671440c8db9d876c** ➡ **Not found**

Today there are many tools that permits generating couple of strings colliding!

```
C:\TEMP> md5sum hello.exe
cdc47d670159eef60916ca03a9d4a007
C:\TEMP> .\hello.exe
Hello, world!

(press enter to quit)
C:\TEMP>
```

```
C:\TEMP> md5sum erase.exe
cdc47d670159eef60916ca03a9d4a007
C:\TEMP> .\erase.exe
This program is evil!!!
Erasing hard
drive...1Gb...2Gb... just
kidding!
Nothing was erased.

(press enter to quit)
```

is possible to create file with the same md5

HASH FUNCTION: SHA

National Security Agency (NSA), 1993 - NIST USA federal standard FIPSPUB180-4.

5 different cryptographic functions

- SHA-1
- SHA-224
- SHA-256
- SHA-384
- SHA-512

} SHA-2

- ❑ SHA-1: digests of 160 bits (most widely used)
- ❑ SHA-2: digests of different lengths SHA-<n> ➡ digest of n bits!

Results	
Original text	*(binary only)*
Original bytes	ffd8ffe2021c4943435f50524f46494c450001010000020c6c... (length=94841)
Adler32	7fb3b6a1
CRC32	9208beae
Haval	cd28f16055cbad8560f06ebb97edf0d4
MD2	98d7dad1c73a0189240683ffa3a8593f
MD4	c2e0c4d449d989ce88c33d056bc8c6ca
MD5	08ee31606c4cedc736c716da0b10329a
RipeMD128	bac154b92a02792Sff93fc3b3ba51d71
RipeMD160	b363eed5bd2f6b079d7177c4ac36db05cb39f1ed
SHA-1	a42c8ba3acf44a35dcb206ee5be88ecf8c946339
SHA-256	d36f2f1c227cab5ee7d03a1ea33faf5f1c0b480315429020bd60012baa1accaa
SHA-384	272964d98b9bb39aea248d0a6167300e16baae6c49403a8a8fd2797751e06587140c3ba7c5e00eac0ac1a024816e08fc
SHA-512	2e89f38aa25cb1f1200655e6d08771d12c415556bb63f9d58b12b2aa566a1ccae790cd44ac56a5a9c5aa80fde686aec237ba0ef2db00bdb4636050d5562cf9d3

CONSERVATION

- Protect data from alterations (natural, on-purpose, damages).
- Crucial whenever evidences are **accessed, transferred analyzed** and.
- Methodology is provided to verify that **no alterations** occurred
- Privacy must be protected as well
- Special packaging could be needed (Magnetic supports, anti-static cases).
- **Cautions:**
    - Data must adequately labelled
    - Verify that batteries are adequately charged (and re-charged), if present
    - Fix mobile parts
    - Minimize risks concerning the type of support
    - Minimize transportation risks
    - Preserve eventual additional traces (Biological traces, fingerprints, Use gloves)
    - Secured in a safe place, protected by a security mechanism (keys), accessed only by authorized personnel,

CHAIN OF CUSTODY (CoC):

- **Definition:** History of the device/support starting from the intervention moment. It is the "Forensic version of log files".
- Documenting material in handling, processing and preserving DE
- Paper/digital documentation
- Crucial in the legal process

CHARACTERISTICS OF CHAIN OF CUSTODY

- **Record:**
    - Who collected DE
    - Where it was stored
    - Who accessed it and why
    - When and how it was transferred
- Must ensure that the DE is untampered and admissible in court
- How?
    - Documentation of collection process
    - Labelling evidences and devices (serial ID), together with hash codes and info-date, time, location of collection;, who collected the evidence and his/her role in the investigation
    - Secured in a safe place, protected by a security mechanism (keys), accessed only by authorized personnel
    - Every access must be logged with 5Ws (who, What, Where, When and Why)
    - Log transfer and handling of evidence

SOME TIPS

- Use proper forensic tools to document and analyze data
- Maintain access logs and Verify and hash data
- Timestamp operations
- All the team must be trained and aligned to specific protocols

ADDITIONAL NOTES ON CHAIN OF CUSTODY

- Document and justify every possible alteration (with the name of the person
- which is responsible for that)
- Record dates if items are released to anyone
- Restrict access to the evidence
- Place original hard drive to a locker
- Perform all forensics on a mirror copy

LECTURE 3

Memory devices are basically everywhere.

MEMORY DEVICE ACQUISITION QUESTIONS:

▪ How many computers/HD to acquire?

▪ What type of computer (e.g., desktop, laptop, server)? (is connected with possible problems with volatility and other stuff)

▪ What are the sizes of the storage for each computer?

▪ Storage device issues (SSD, Encryption, …)? (SSD have a problem and is inherent problem in the SSD, will see later)

DATA STORED ON DIGITAL DEVICES

Problem: retrieve relevant information from a memory device (even information erase or remove).

- Number of possible files and application is very large
- Look only for what is relevant to the specific case (but this could require an exhaustive analysis, anyway)
- Disk and operating system investigation tools exist in the commercial, (you prefer) use free and open-source domains.
- Operations must be accurately chosen depending on the device and the data to be retrieved

Given that we need to pose 2 questions:

- What to look for? Many traces can be found
- How to look for?

We respond to this question watching how the hard disk is organized.

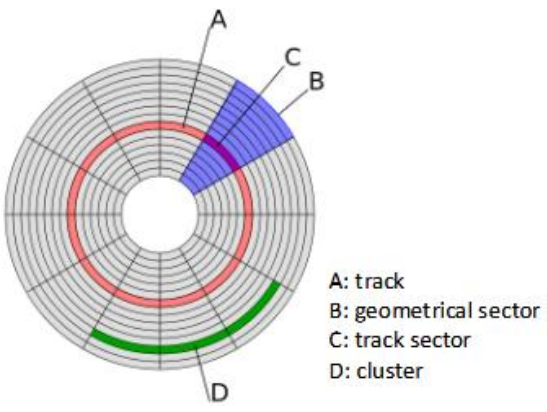
MEMORY SPACE ON A DEVICE

In computer file systems, a **cluster** or **allocation unit** is a unit of disk space allocation for files and directories. To optimize disk management, filesystems do not allocate individual sectors but contiguous groups of sectors (clusters).

Allocated: used to store data, metadata, programs

Non-allocated: available

Slack space: area between the last bit of a file and the end of the last assigned sector (wasted). Sometimes the term is extended referring to any part of a memory which is not indexed any more.

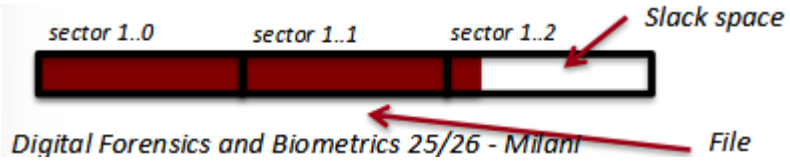


Every sector can be assigned to a single file; a single file can be stored in multiple sectors.

Generally, the information is stored into cluster (D), we want to quantize the information you want to write.

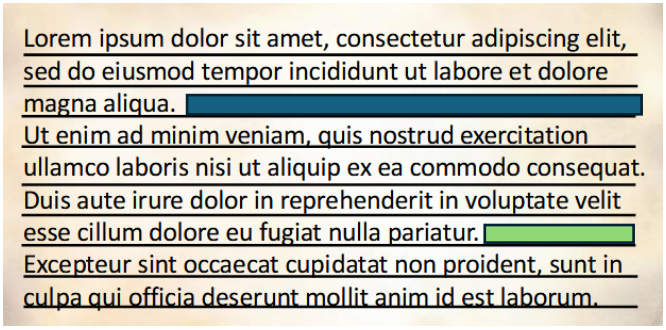
We don't allocate in a single bit but we allocate in sectors (basic unit of memory).

Definition of Allocate: space/part of the disk occupied by a file.



Each sector can store 8Kb but I have 9Kb so I need 3 sector for store this file. There are 3Kb left not use -> this is the slack-space. Sometime slack-space can be use to define other types other objects, for example I have a space occupied by a file which was removed and in that case that space is consider slack because the information is still there but is considered removed.

EXAMPLE



Each raw is considered as a cluster, so the first sentence occupied 3 cluster and the ■ is the slack space.

The second sentence are 4 cluster with ■ slack-space.

TRACES FROM DISK SPACES

When you remove a file the operating system things that now it can use that space. Files are indexed by the filesystem (one or more per file). Removing the file implies deleting the indexes: data are still present on device until they are overwritten by another file. Infos are there.

For example we deleted the File A, this means that this sector are free but the bits of File A are still there (you removed only the reference).

Now we write File B, part of the data of file A are still present!

Note that before a space is overwritten, it would require a long time.

Some data can be retrieved from the so-called **G-list**: damaged sector that can not be used any more but that can be read!

45 % of data that can be read from a hard disk are in slack and non-index spaces!

SOLID STATE DRIVE

SSD works differently compared to hard disk. The file recovering can be challenging.

• SD store data on flash memory chips.

• No delays due to spinning disk (simultaneous access)

• File recovery is much more challenging!

Basic problem in SSD: blocks need to be erased before being written. For this reason the SSD have 2 commands:

1. **TRIM** (TRIM in ATA set or UNMAP in SCSI)

   • issued by OS to SSD controller at file deletion, disk formatting, or partition removal

   • Basically this command says which block of data are no longer used

2. **Garbage Collection**

   • cleans or purges the data blocks marked as deleted.

   • Implemented on disk itself (not controlled by OS): if PC is shut down, it restarts when powered on or even if attached to a write blocker, when PC restarts this function can be automatically activated.

Another problems: Encryption and Compression

Result: purging of data block is unavoidable! But still we can do something…

Data blocks and file carving is still possible anyway

TRIM limitations:

• TRIM is not supported sometimes,

• PC support RAID environment, external SSD and NAS.

• SSD connected through USB (TRIM is only supported in SATA, eSATA, and SCSI)

• buggy firmware in the SSD.

• the file is just corrupted and not deleted.

Example: SATA protocol, different protocols after TRIM command

1.Non-deterministic TRIM: each read command may return different data

2. Deterministic Read After TRIM (DRAT): all read commands shall return the same data,

3. Deterministic Zero After TRIM (DZAT): all read commands shall return zeroes until the page is written with new data.

Now day the majority of computer us SSD.

WHAT TO DO ?

• Attach SSD to a secondary channel of the motherboard (not primary drive)

• Check if disk is close to full. The closer is to being full the higher is the probability that is your data is been overwritten to other data.

• Check the garbage collector characteristics (dipend on vendor related issue)

FILESYSTEMS

Filesystems handle data in a different way: we need different strategies!

In filesystem Data are fragmented into multiple (and sometimes not contiguous) sectors: indexing information permits retrieving all the parts. This fragmentation is possible because the index.

Removal:

- Erase the address of the table that contains the indexes of all the parts of the file
- Files are moved to another directory (trashbin)

Data are still there unless is

- overwritten
- removed with wiping software

DATA CARVING

The data do not have the index anymore but are not overwritten and you want extract them. Files can have different extensions so we look in specific format. (Depends on the type of file)

Magic number is a constant that identifies the file type

PDF: starts with %PDF and ends with %EOF

JPG: starts with 0xFFD8 and ends with 0xFFD9

(use hexedit or hexdump to view; file on Linux) so if I look in a part of hard disk and I see *0xFFD8* I suspect that here starts a jpeg file.

- **Data carving** is a process of extracting a set of data/information from a wider set of data
- In forensic analysis data carving is used to find data in the non-allocated space
- File classified from headers and footers
- Structure of filesystem is ignored.

WHAT DATA CARVING ALLOWS

Base conditions:

• Header and footer have not been overwritten

• The file is not fragmented (not too much fragmented)

• The file is not compressed (sometimes work with compressed/corrupt files)

• The file is included between header and footer

Advanced conditions:

- Non-sequential fragments
- Non-ordered fragments
- Some fragments are missing

❑ Recovering deleted or fragmented files (fragments) from damaged memories; it can be used on RAM dumps as well.

❑ Performed on device or image

❑ Footer can be missing

❑ Beware of false positives!

Tools: *foremost, scalpel, photorec* (they work on blocks). If you will use this tools for recover some data that you have lost, with this data you will recover a lot of other information that you even know you have, this is due to the fact we will have a lot of false positives.

FOREMOST

The problem is that your not recover file names (audio.mp3) but you recover strings/sector address.

```
Foremost version 1.5.7 by Jesse Kornblum, Kris Kendall,
and Nick MikusAudit
FileForemost started at Fri Apr 20 10:15:45 2011
Invocation: foremost /dev/sda1
Output directory: /data-recovery
Configuration file: /usr/local/etc/foremost.conf
-------------------------------------------------
File: /dev/sda1
Start: Fri Apr 20 10:15:45 2011
Length: 962 MB (1009254400 bytes)
Num Name (bs=512)
Size File Offset Comment0:
0: 00000621.jpg 11 KB 312832
1: 00000743.jpg 8 KB 329216
2: 00000774.jpg 3 KB 345600
3: 00000877.jpg 9 KB 361984
4: 00000983.jpg 8 KB 378768
```

Byte size of sector is 512

It was on sector 611

Offset is 512*621=312832
(byte address of file on disk)

Verify sector on disk

```
$ xxd -s 312832 -l 512 /dev/sda1

004c600: ffd8 ffe0 0010 4a46 4946 0001 0100 0001 ......JFIF......
004c610: 0001 0000 ffdb 0043 0005 0304 0404 0305 .......C........
004c620: 0404 0405 0505 0607 0c08 0707 0707 0f0b ................
004c630: 0b09 0c11 0f12 1211 0f11 1113 161c 1713 ................
004c640: 141a 1511 1118 2118 1a1d 1d1f 1f1f 1317 ......!.........
004c650: 2224 221e 241c 1e1f 1eff db00 4301 0505 "$".$.......C...
004c660: 0507 0607 0e08 080e 1e14 1114 1e1e 1e1e ................
```

PHOTOREC

Photorec is another software (an open source one) for recovering files. Recovery tool for many types of files. Can be used together with TestDisk. It recognize 400 different formats (ZIP, Office, PDF, HTML, JPG, etc...)

AUTOPSY + SLEUTH KIT

Sleuth Kit is a suite of tools for disk and file system forensics. Autopsy Forensic Browser is its interface

• Permits analyzing, disks, images,

• Generate a timeline

• Generate hashes

• Permits looking for file types

• Permits looking for a specific content in files

• Looks for erased files

DISK ENCRYPTION

Tree type of encryption:

SW (Software) encryption is:

- run by an application
- slower
- tied to a set of keys but not bound to a specific machine or platform. So one time I have the keys I can access to the ecrypted data from any devices
- E.g. TrueCrypt

As for **HW (Hardware) encryption**:

- operated at firmware level; relies on Trusted Machine Platform (TPM) chip
- entire drive (incl. boot sector, OS, temporary files, and swap files) are encrypted
- Faster, managed by the control system of the HD

Hybrid encryption instead:

- is software but can be linked to a specific HD
- relies n TPM as well
- access can be in transparent operating mode, user authenticated or USB authenticated
- eg. Windows BitLocker

HOW TO CIRCUMVENT?

If you have SW encryption you could:

- Cold reboot attack (exploits DRAM not completely erased - boot a malicious operating system that copies the contents of memory)
- Pieces of files can be found outside the encrypted area (temporary storage) -

  Correct and incorrect password – it depends how encryption is married with OS

HW encryption

- Pieces of files can be found outside the encrypted area
- Avoid sleeps or shutting down: when re-enacted password must be entered
- Some systems aim at creating recovery keys: stored in clear view, they can be used to decrypt the disk
- Exploits TPM weaknesses: keys can be stored in clear

Anyway, encryption can be extremely useful in destroying evidences: reset/wipe encryption keys

MOBILE DEVICES

Recovering traces from mobile devices involve more than simply files. This is a problem because you have a lots information compared to pc.

Traces: IMEI, SIM cards, images, ...

Steps:

- Identification and evidence collection
- Acquisition and verification
- Conservation
- Valuation and presentation

IDS FOR MOBILE DEVICES

IMEI codes: every mobile terminals is identified by a 15 characters code

(International Mobile Equipment Identifier or IMEI), used to identify the device within the mobile network. It can identify:

- Vendor
- Model
- Nation

How to get it?

Switched off devices: look on the back or under the battery

Switched on devices: dial ***#06#**

Subscriber Identity Module (SIM) card: required to access a mobile network.

Every SIM card is characterized by a Integrated Circuit Card Identification (ICCID) number, it's like an ID for the SIM, it contains encoded information about the network provider and the country.

Tool: http://www.numberingplans.com/ allows to get information about the network given the ICCID code.

When a Digital Forensics expert arrives at a scene, they must follow a strict protocol to ensure evidence is valid in court:

1. Secure the phone/tablet
2. Isolate it and prevent other people from operating with it
3. Notate eventual physical defects (e.g., broken display, etc.)
4. Take pictures of the device
5. Document all the actions
6. Verify if switched on or off
7. If the device is off, leave it switched off. Turning it on might trigger boot sequences, alter files, or connect to a network that could send a "remote wipe" command.

In case the phone is on,

- Document the information on the screen
- If possible, register date and time of the device verifying the eventual offset with respect to the real time.
- Avoid navigating the menu or opening messages
- Keep it switch on isolating it from the possible network connections
    - Bluetooth 2,45GHz
    - Wi-Fi
    - GSM/UMTS/LTE
    - GPS1575MHze1227MHz

  Isolation can be done with:

    - **Jammer**
    - **Faraday's cage** (aluminum foil) (better than Jammer)
    - **Airplane mode** (verify all connection are off)

or

- Switch it off with normal shutdown or removing the battery

SWITCHING OFF VS. ISOLATION

Switching the device off could enable phone lock and require an unlocking code (e.g., PIN of SIM card or lock number of the phone).

Using jammers or Faraday's cages imply a waste of battery (the phone keeps on trying to connect) need for external power charge to keep it on.

Airplane mode isolates the phone without risks … anyway, if you do not know the model you still can mess up the things.

OTHER STEPS IN SEIZURE

Other things to collect with the phone:

- Cables
- Battery charger
- Wrappings
- Memories
- Manuals
- CDs, USB pens storing software suites to manage the phone
- Phone bills related to the phone
- SIM envelope with PIN, PUK codes

Document the seizure with the following data:

- Operator's name
- Date and time of the seizure
- Where the device has been taken (address, GPS position, etc.)

DATA ACQUISITION ON MOBILE DEVICES

Data storage on a mobile devices are:
(different part where the information can be memorize)

- ❑ SIM card
- ❑ SD card
- ❑ Integrated memory
  - Logic acquisition (files)
  - Physical acquisition: cloning NAND memory

SIM card: see next slides, include information about the connection mechanism

SD card: removable devices, strategies are similar to those seen for HDs

Integrated Memory: Not only the memory present on the phone but also the memory present in the "registry". Logic acquisition can be done using:

- softwares (specific for forensic analysis or for data extraction), connection to access to the phone. Need for a reliable connection between mobile device and computer such as:
  - Cables 🙂
  - Infrared 😐
  - Bluetooth (problem: change the state of device during the connection) ☹
- hardware devices.

Physical acquisition tools

- depend on the specific device and OS
- permit recovering lost data

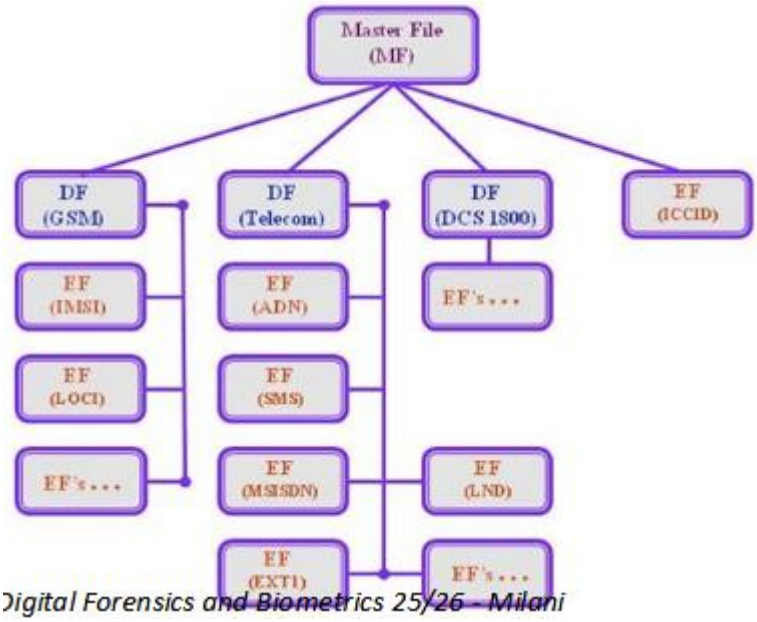
DATA ACQUISITION FROM SIM CARDS

SIMs can be locked and data can be often cyphered using passcode numbers.
Need to unlock the phone

- ❑ PIN code (Personal Identification Number) – 4 digits - 3 trials
- ❑ In case of 3 failures, you need to use PUK (Personal Unblocking Key) – 10 digits – 10 trials

If you've arrived here, your SIM needs to be replaced

SIM memory has a tree structure. The tree structure is constituted of folders.



Folders **DF(GSM)** and **DF (DCS1800)** contains information about the network, while **DF(Telecom)** contains data regarding the activated services.

Data that can be retrieved

- ICCID(Integrated Circuit Card Identification)
- IMSI(International Mobile Subscriber Identity)
- Numbers (Abbreviated Dialing Numbers – ADN)
- Call records(Last Dialed Number – LDN)
- Short Message Service (SMS)
- Short Message Parameters (SMSP)
- Location information (LOCI)
- SIM Service Table (SST)
- Public Land Mobile Network (PLMN) selector
- Forbidden PLMNs
- Service Dialing Numbers (SDNs)

HOW TO READ SIMS?

❑ Use a SIM reader

❑ Support standard PC/SC (http://www.pcscworkgroup.com/)

Main softwares to read a SIM card

- SIMiFOR - http://www.forensicts.co.uk/ (commerciale)
- SIMcon - http://www.simcon.no/ (commerciale)
- USIM Detective - http://www.quantaq.com (commerciale)
- DekartSIM Manager - http://www.dekart.com (commerciale)
- SIMSpy2 - http://www.nobbi.com/ (freeware)
- Tulp2G - http://tulp2g.sourceforge.net/ (freeware)

MAIN SOFTWARE TOOLS FOR INTEGRATED MEMORY ACQUISITION

Software for backups, for cloning (iTunes, BlackBerry Desktop Manager, Nokia Suite, Samsung Kies)

Forensic software to acquire information on the phone(Oxygen Forensics Suite, Compleson Lab MOBILedit! Forensic, Paraben Device Seizure, Mobile Phone Examiner)

Hardware solution: the problem of HD solution is that it's very limited by the specific version of the phone. (Cellbrite UFED, Micro Systemation XRY, CellDEK)

LECTURE 4

INTRODUCTION TO NETWORK FORENSICS

Network forensics: monitoring and analyzing computer network traffic to gather information, compile evidence, and/or detect intrusions. The difference compared to the Lecture 3 is that in the last lecture the evidences are static here the evidences are changing over time.
Network forensics are 2 main tasks to achieve:

- Volatile, because data are changing over time.
- Rarely-logged, sometimes you are logs the information but not all information are log. (there is sometimes encryption)

The data that we acquiring somehow are evolving, so we are not interested in a single information but more how this information is changing over time. So now we don't have no mre an image or files but we are measuring the "time-c".

Uses and application of network forensics:

Security

- Monitoring for intrusions, using detection system
- Network evidence may be the only type if a drive was wiped clean

Law enforcement

- Reassembling transferred files
- Finding keywords
- Searching for keywords
- Parsing messages
- Examining packet filters
- Examining firewalls
- Examining existing systems

NETWORK FORENSICS OPERATIONS

Multi-sensors data fusion

An attack often is not isolate, for example an attack is sensed/registered from multiple machine or multiple elements. This element are different and often you have to combine them.
Collect the logs from all network security products, deployed in the entire network and perform data fusion.

- packet analyzers (wireshark / tcpdump)
- intrusion detection systems (snort),
- routers, firewalls, log servers, etc.
- alerts generated by IDS,
- statistics from protocol analyzers and attack information by observing various threshold values

Dempster-Shafer theory for information fusion determines the validity of the attack.

Identification of Attack Events

Large amount of memory and storage is usually required; network events useful for investigative requirements need to be identified and an effective mechanism is to be in place to identify attack features from the traces. For the identification will use specific pattern to identify what type of attack.

Attack Reconstruction

Important events involving intruders' interaction with the compromised system are reconstructed and the methodology of the attackers is analyzed. So we want to know how this attack was place.

Traceback and Attribution

Determining the origin of a packet. Techniques based on packet marking, packet logging or hybrid approaches. Attribution relies on analyzing the data packets transmitted, applications being run, traffic patterns observed and protocols violated.

Incident Response

Are to be launched immediately when the alerts begin; attacker must not be aware of the response. Detection and validating the incident by reviewing pertinent logs, network topology, etc. It determines the vulnerability exploited in the compromise of a system and enforces protection against exploitation of the same on other systems. It develops a strategy regarding containment, eradication, recovery, and investigation.

NETWORK FORENSICS' STEPS

Identification: recognizing and determining an incident based on network indicators.

Preservation: start on the very beginning, securing and isolating the state of physical and logical evidences from being altered, such as, for example, protection from electromagnetic damage or interference.

Collection: Recording the physical scene and duplicating digital evidence using standardized methods and procedures.

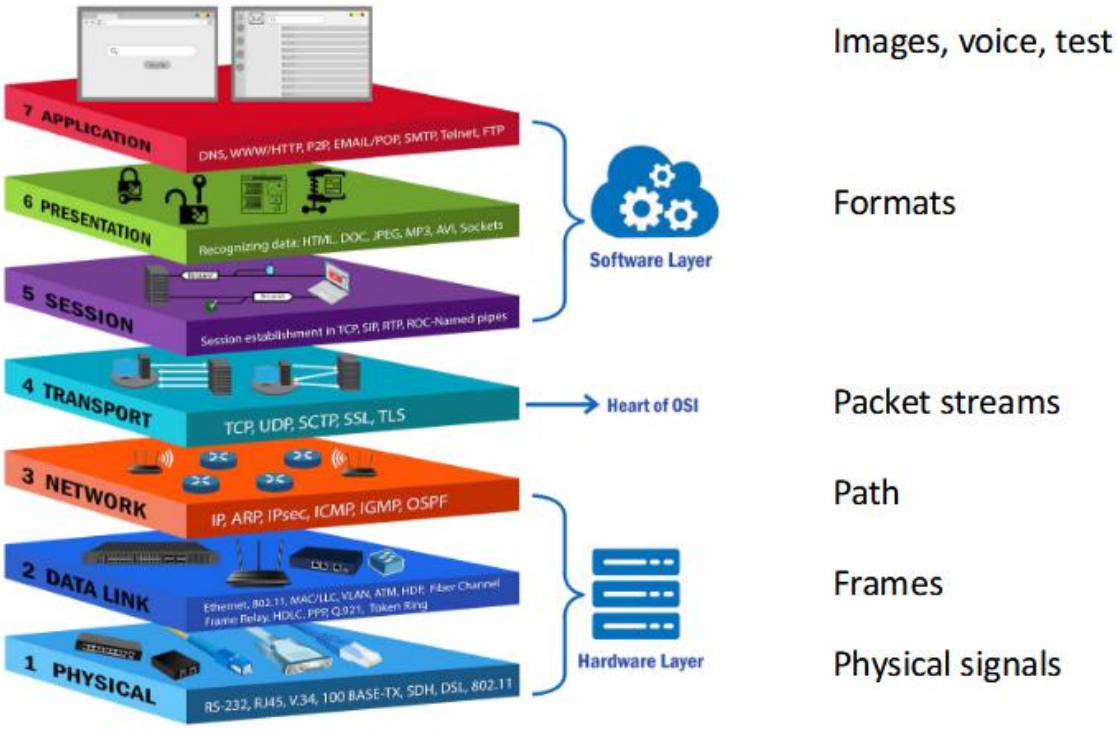
Examination: in-depth systematic search of evidence relating to the network attack. This focuses on identifying and discovering potential evidence and building detailed documentation for analysis.

Analysis: determine significance, reconstruct packets of network traffic data and draw conclusions based on evidence found.

Presentation: summarize and provide explanation of drawn conclusions.

Incident Response: The response to attack or intrusion detected is initiated based on the information gathered to validate and assess the incident.

OSI STACK STRUCTURE



Osi stack is a theoretical model:

7. **Application Layer:** Interacts with user software and provides network services (e.g., HTTP, FTP).
6. **Presentation Layer:** Handles data formatting, encryption, and compression.
5. **Session Layer:** Manages the setup and termination of connections (sessions) between applications.
4. **Transport Layer:** All the transfer protocols are here. Ensures reliable data transfer end-to-end (e.g., TCP/UDP).
3. **Network Layer:** Deals with logical addressing (IP) and routing across different networks.
2. **Data Link Layer:** Handles physical addressing (MAC) and data transfer over a single network link.
1. **Physical Layer:** Transmits raw data bits over the physical medium (cables, interfaces, radio wave).

OPERATING LAYERS IN NETWORK FORENSICS

Depending on the layer network forensics do different task:

- **Ethernet analysis** (Data Link layer)
  - Process packets on the same network segment
  - permits profiling activities and analyzing local behaviors (alarms)
- **TCP/IP – Router analysis** (search how the data is transmitted to the network)
  - Mainly involves data tracking
  - Follow compromised packets, reverse route, ID the source
  - Network layer also provides authentication log evidence

- **Internet/Server based analysis** (upper layer of OSI stack)

- Includes web-browsing, email, chat, and other types of traffic & communication
- Server logs collect information
- Email accounts have useful information except when email headers are faked
- User account information associated with a particular user

➢ **Wireless analysis** (the radio part, similar with ethernet forensics, the problem are that here you have to deal with wireless channel)

- A sub-discipline of the field
- To get that which is considered "valid digital evidence"
- This can be normal data OR voice communications via VoIP
- Analysis is similar to wired network situations, with different security issues

DATA COLLECTION METHODS

"Catch-it-as-you-can"	**"Stop, look, and listen"**
• All packets/data are captured • Large storage needed • Analysis in batch mode • Usually @ packet level • For later analysis • All the packets are capture • Problem: a lot of data to analyze	• Requires faster processor for incoming traffic • Each analyzed in memory • Certain ones are stored • Usually @ packet level • Real-time filtering • You search accord to a precise criteria a specific elements • Problem: you need analyze and then acquire, so is very slow

See packet sniffers or analyzers: "sniffer captures each packet and, if needed, decodes the packet's raw data, showing the values of various fields in the packet, and analyzes its content according to the appropriate RFC or other specifications".

Example of the Lab: we collect the packet with WireShark, analyze the length of the packet, because the packet are encrypted so you can't access to the information. Using only the length of the packet was possible to understand if the user was using telegram, or Instagram etc.
You can use this with video surveillance streaming and track a person with the length of the signal.

ETHERNET

Technology used in used in local area networks (LAN), metropolitan area networks (MAN) and wide area networks (WAN)

❑ High bit rates 2.94 Mbit/s ('76) to 400 Gbit/s (today)

❑ Data-link layer

❑ Data stream partitioned into frames

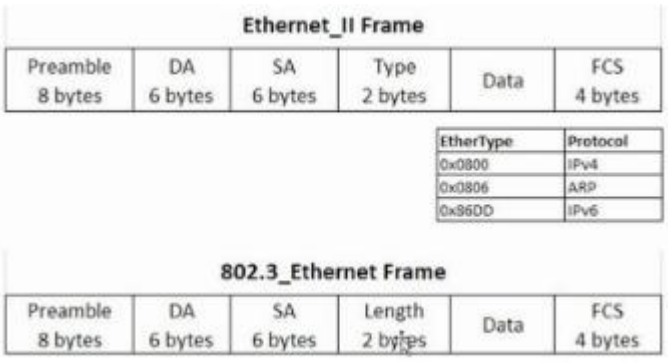
- Source
- Destination

} 48-bit MAC address

- Error checking code

DA: destination address
SA: source address

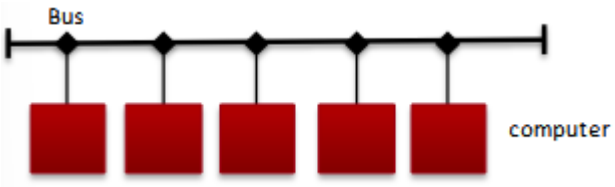
Structure of a packet:

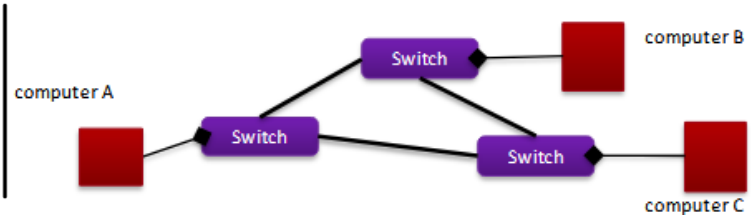


How the packet is transmitted throw the network? All the computer look at the packet once is transmitted, if a computer saw that the DA is his DA it acquire it.

Sniffer: it's acquired/read all the packets regardless of the DA. Work in *promiscuous mode*, is a mode of the ethernet card that forces to acquire all the packets.

Shared Media: any information is received by all; interrupts happen only when applicable packets are received: the card ignores information not addressed to it.





Original Ethernet has problem:

- Collision
- Low throughput

Modern Ethernet solve that problem introducing switch (they guide the packets):

- Station -> switch -> destination (lower collision probability)
- 10-Base T: full duplex mode (collision free

Employed devices:

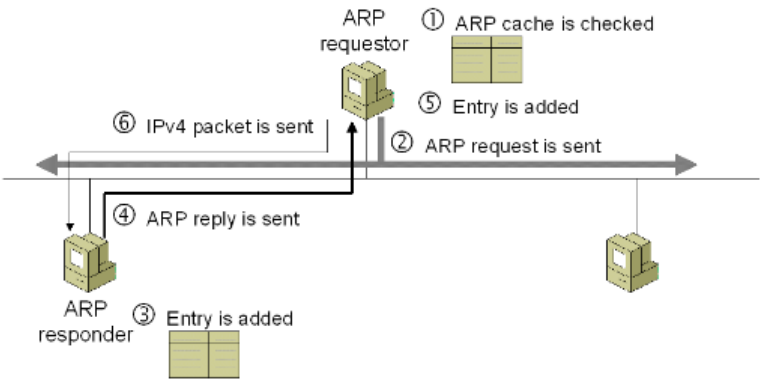
- repeater and hubs (cope with signal degradation and timing reasons; forwards to all)
- bridges and switches (selective forwarding; frail to single points of failure, dumping attacks that trick switches into sending data to a sniffing machine, scalability and security issues e.g. switching loops, broadcast radiation and multicast traffic, bandwidth choke points).

HOW SNIFFER WORK

Every PC on a LAN has two addresses: MAC (used by card) and IP (used by protocols)

ARP: Address Resolution Protocol (RFC 826) maps IP address into MAC addresses

ARP cache stores all the pairings IP/MAC addresses. You basically exploit a specific protocol which is used to assign a specific IP addresses. A PC has two main addresses: an **IP address**, which is a network address for sending data, and a **MAC address**, which is a physical address that uniquely identifies the hardware on a local network. The IP address is like a street address that can change depending on the network, while the MAC address is like a serial number permanently assigned to the network card by the manufacturer.



In ARP you have a memory table (**ARP cache**) and for each entry you have an association between IP address and MAC address. ARP works by using a two-step process: a device sends a broadcast ARP request to find the MAC address associated with a target IP address on a local network, and the device with that IP address responds with an ARP reply containing its MAC address. The requesting device then stores this IP-to-MAC address mapping in its table for future communication, avoiding the need to send another broadcast for the same device.

Sniffer on shared Ethernet (passive)
A machine running a sniffer accepts all the frames regardless of the MAC address (promiscuous mode)

Sniffer on switched Ethernet (active)
Previous strategy does not work.

ARP spoofing: sends an ARP (fake) reply even if it was not asked for. It replace the correct MAC address with its own. All the traffic passes Through the eavesdropper. (digitally identity theft)
Linux command: arpspoof

MAC flooding: switches keep a table mapping MAC addresses to physical ports: limited memory for this. Flod the switch with false MACs until the switch cannot keep up: *failopen mode*. In this mode the switch then operates like a hub broadcasting sending all the packets to all the machines.
n.b. This method degenerates network services – must be used for a limited amount of time!
Linux command: macof

EAVESDROPPING PACKET STREAM
Can do for malicious scopes or for protection.

- Uses monitoring tools or sniffers (traditional Eth: sniffing machine; modern Eth: switch with monitoring port)
- Wireshark(a.k.a. Ethereal)
- Then protocols can be consulted, such as the Address Resolution Protocol (ARP)
- Network Interface Card (NIC), but can be averted with encryption

Uses of sniffers:

- Analyze network problems
- Detect network intrusion attempts
- Detect network misuse by internal and external users
- Documenting regulatory compliance through logging all perimeter and endpoint traffic
- Gain information for effecting a network intrusion
- Isolate exploited systems
- Monitor WAN bandwidth utilization
- Monitor network usage (including internal and external users and systems)
- Monitor data-in-motion
- Monitor WAN and endpoint security status
- Gather and report network statistics
- Filter suspect content from network traffic
- Serve as primary data source for day-to-day network monitoring and management
- Spy on other network users and collect sensitive information such as login details or users cookies (depending on any content encryption methods that may be in use)
- Reverse engineer proprietary protocols used over the network
- Debug client/server communications
- Debug network protocol implementations
- Verify adds, moves and changes
- Verify internal control system effectiveness (firewalls, access control, Web filter, spam filter, proxy)

LECTURE 5

Sniffer: is a tool used to intercept, log, and analyze data packets as they travel across a network.

HOW TO DETECT A SNIFFER

Ping method
Send a packet with the IP address of the suspect machine without or with a wrong MAC address. All machines should reject it since MAC does not match. Only the sniffer answers (it does not bother accepting packets with a different MAC address).

ARP method
Exploits ARP cache. Sends a non-broadcast ARP (contain information about the association between IP and MAC), which is cached by a machine in promiscuous mode (this packet will contain wrong matching between IP and MAC). Then, send broadcast IP packet with correct IP but different MAC address. Only machines with correct MAC address from the sniffed ARP frame will respond.

Local host analysis
Hackers may have compromised your terminal and left sniffers. Use *ifconfig* and analyze the answer. Search for strange activity on the device.

Latency method
Often sniffers are acquiring a lot of traffic so these machines are typically slow. Most sniffers do some parsing. Huge amount of data is sent through the network, and during the transmission, the suspect machine can be **pinged**. In promiscuous mode, it parses all packets and therefore, ping response is delayed. False positives are possible (traffic can be delayed by the network load). This method differs from the Ping method because here we exploit the timing in the other one we exploit the wrong matching.

ARP watch
Especially when you have a man-in-the-middle attack (your ARP cache is compromise), you can verify the changes of the ARP cache you can detect suspect behavior. Gateway can be spoofed. *arpwatch* can be used to check the ARP cache of a terminal: look for duplicates. With DHCP there can be many false alarms: increase the lDHCP lease time.

IDS
Watch traffic. Snort record IP/MAC pairings of packets. Whenever a mismatch is found, it generates an alert.

SNIFFING SOFTWARES

- **Tcpdump**: works on Linux systems, old software (sniffers' granddaddy), permits real-time filtering). www.tcpdump.org/daily/tcpdump-current.tar.gz
- **Wireshark**: sniffs at both frame and packet level, allows filtering. https://www.wireshark.org/
- **Snort**: sniffs and do packet logging as well, can be used for IDS. https://www.snort.org/
- **Ethereal**: Similar to previous ones, live captures. https://www.snort.org/
- **Hunt**: Exploits weakness in TCP/IP protocol; allows hijacking active connections and take over their control.

- **Ettercap**: Designed for switched LAN; allows performing MITM attach against SSH and SSL; performs password collecting for several protocols. http://www.ettercap-project.org/
- **Dsniff**: Comprehensive sniffer package. https://www.monkey.org/~dugsong/dsniff/

ANTI-SNIFFING MEASURES

- Tools that protect the ARP against Man-in-the-Middle (MITM) attacks

  e.g. ARP Spoofing, ARP Cache Poisoning, ARP Poison Routing an attacker sends (spoofed) ARP on LAN. The aim is to associate the attacker's MAC address with the IP address of another host, such as the default gateway.

- Cryptography (SSH, SSL, https, PGP or GnuPG for emails
- Sniffing detection tools
  - **Snort**: enables IDS
  - **Anti Sniff**: detects if a computer is in promiscuous mode
  - **ARP watch**: keep track of MAC/IP address pairings
  - **Neped**: (outdated)
- Stronger security measures on network protocols
  - implemented on switches, gateway, etc. they perform a set of checks in real-time, like whether a MAC address I is associated to 2 IP addresses, and so on.
  - add MAC address permanently for authorized machine in your networks.

PCAP FILES

Stores information about sniffed streams (tools: libpcap)

Global header: which defines information about the stream you'll analyze. This information are:

- defines format and ordering,
- correction (in sec) between time zone and GMT (UTC),
- accuracy of time stamp,
- "snapshot length" for the capture,
- link-layer header type.

Packet header: include information about the packet:

- timestamp (in sec),
- timestamp (in microsec),
- Length of stored bytes from packet,
- Length (real) of the transmitted packet.

Packet Data: the actual data of the packet.

TROJAN HORSE EAVESDROPPING

SMS Trojan, backdoor trojan. Trojan differently from sniffer which focuses on the streams, here are full control of the device. In case the infected machine have a exchange of packets with another machine you can force the packets throw your machine. But you can do a lot more because you have the full control of the machine. This is different from the MIMA because this forces the network to send the

packet to your machine and the your machine will re-route the packet, here the correct destination "work" for you.

- Sniffer exploits vulnerability of network streams
- It is possible to take control of devices
- Use of Trojan Spyware
- Collect communications, logs, voice call, images/videos (i.e., computer and sensed activity).
- Different from MITM (attacker is a relay)

How get control? Click on link, SMS, phone call, install malicious app

For legal reasons the company that give this type of service don't develop the software themselves, but they take from third parties. They simply collect differ software for different devices. Another problem is that the malware compromises the system.

ROUTING ANALYSIS

TCP/IP ROUTER

- Device specialized on sending packets across data networks
- Responsible for Interconnecting networks by selecting the best path for a packet to its destination
- Forward packets from one network/segment to another one (according to a routing table)
- Different routes are possible
- Information is stored on routing tables; router decide to send a packet to a specific route based on the routing table.

ROUTING ON TCP/IP

Routers communicating throw:

- Local cable
- Modem (in case the router is too far and it can not use a cable)
- Terminal emulation software

Components:	Ports:	Working modes:
▪ ROM ▪ POST ▪ IOS ▪ RAM ▪ Flash memory ▪ NVRAM	▪ LAN Ports ▪ WAN Ports ▪ Administrative ports - console ports - auxiliary ports	▪ Setup mode ▪ User mode ▪ Privileged mode ▪ Global configuration mode ▪ Interface mode

ROUTER ATTACKS

Importance of preserving routers from attacks. Router processes a large amount of data so are must be very quick. Updating the firmware is important for security, you can do locally or from remote.

- Heart of networks
- Few procedures for hardening security
- Slower in getting upgraded
- Few people monitor their configuration regularly
- Few security measures
- There are billions of them

Attacks:

- Reconnaissance
- Scanning and enumeration
- Gaining access
- Escalation of privilege
- Maintaining access
- Covering tracks and placing backdoors
- DoS attacks
- Packet mistreating attacks
- Routing table poisoning
- Persistent attacks

MOST COMMON ROUTER ATTACKS

- **Denials of Service (DOS)**
  Used by hackers to disrupt an entire network and router. Series of requests; flood the routers networks with message requests, e.g., Internet Control Message Protocol (ICMP) packets; sent over a short time from multiple locations (in this way is more difficult to track you).
- **Packet mistreating**
  Injects packets with malicious codes designed to confuse and disrupt the router and network. As the routers become more and more confused, malicious data starts to circulate around the network creating a loop = congestion
- **Routing table poisoning**
  Very similar to ARP poisoning. These aggressive attacks are achieved by editing the information packets that are cycled through the routing table.
- **Hit and run attacks**
  One off attack on a specific network or router: malicious data is injected into router through code. If attacker fails at their first attempt, they may or may not progress and make further attempts on the system. Attack not really sophisticated but they relay on the speed.
- **Persistent attacks**
  Occur and continue to occur until the attacker has achieved their goal.

GATHERING VOLATILE ROUTER DATA

If you want to gather information from the router you:

- connect to console port; cable (it's better to take them from the cable if it's possible) + laptop with terminal emulation software
- record System Time and determine who is logged on
- save router configuration
- review routing tables and detect malicious static routes modified by attacker view ARP cache looking for evidence of IP and MAC spoofing

Order of volatility you need to follow:

1. Registers and cache.
2. Routing tables.
3. Address Resolution Protocol cache.
4. Process table.
5. Kernel statistics and modules.

6. Main memory.
7. Temporary file systems.
8. Secondary memory.
9. Router configuration.
10. Network topology.

This works if you connect to the router from a cable. If you can't connect using a cable you have to connect remotely but it's riskier.

CAUTIONS AND DOCUMENTING

DOs

▪ Access router through the console
▪ Record all your console session
▪ Run show commands
▪ Record actual time and router time
▪ Record volatile information

DON'Ts

▪ Reboot the router (you lost precius information)
▪ Access the router through the network (if it possible avoid that)
▪ Run configuration commands
▪ Rely only on persistent information

Documentation

- *Chain of custody*: prove the integrity of evidences.
- *Case reports*: employee remediation, employee termination, civil proceedings, criminal prosecution, case summary, bookmarks
- *Incident response*: effort to define and document the nature and the scope of a computer security incident

BAD GUYS AND GOOD GUYS

▪ *Internet Router Protocol Attack (IRPA) suite*: a suite of tools designed to abuse inherent design insecurity. tools: ass, igrp, hsrp
▪ *VIPPR*: can be used for MITM attacks for compromised routers
▪ *UltimaRatio*: tools that exploit vulnerabilities on routers
▪ *Research*

In other to defend yourself you can:
▪ *Router Audit Tool (RAT):* Scores overall security on router
▪ *Manuals, books, reviews, papers*: on securing routers
▪*Strong authentication*: Encrypted traffic mgmt, two-phase authentication, centralized authentication source

SERVER LOG

Can verify the information about connection and network traffic. We are focuses first in server Logs.

- Are automatically generated and maintained by a server, it records all activities and events that occurred. (it's like a diary)
- detailed record, like a diary, of everything the server does, including access attempts, data requests, and any errors encountered

Typically the Logs are use to verify the correct function about a service.

Server logs are useful for:

- **Troubleshooting**: pinpoint the cause of errors and identify unusual behavior on the server.
- **Security**: reveal suspicious activity, like multiple failed login attempts or unusual access patterns

- **Performance monitoring**: insights into website traffic, resource usage, and potential bottlenecks.
- **Compliance**: auditing purposes to ensure compliance with regulations and data protection laws.

- You can have different types of logs. Formats include plain text, structured, binary, syslog, CLF, and ELF. This can be generated by Antivirus program, firewall, intrusion detection/prevention system, vulnerability management, authentication server, router

Log Type	Description	Common Contents
System Logs	OS events, processes, and errors	• Kernel logs • Authentication Logs • Security Logs
Application Logs	Capturing events and errors by applications	• Contextual Information • Timestamps • Log Levels
Security Logs*	Logs tracking security events and alerts	• Authentication logs • Access logs • Intrusion detection
Audit Logs	Logs used for compliance	• User actions • Data access events • System changes
Event Logs	Logs tracking significant system, security, and application events	• System events • Application-specific events

COMMON LOGS FORMAT

NCSA Common log format is a common format, each log is a entry that stores information about: [host; ident; authuser; date; request; status; bytes]

Example: 127.0.0.1 user-identifier tomasin [30/Aug/2017:10:25:16 -0700] "GET /apache_pb.gif HTTP/1.0" 200 1068

Extended log format (ELF) with this information: [host; ident; authuser; date; request; status; bytes]

Logs play an important role in tracking each client computer's activity

- keep an eye on your network for vulnerabilities
- identify who introduces risks, and help that person to use better precautions.

Same approach can be applied to any log file (application, machines, ...)

WIRELESS CHANNEL SENSING

INTRODUCTION TO WIRELESS FORENSICS

methodology and tools required to collect and analyze (wireless) network traffic that can be presented as valid digital evidence.

How did they distinguish from wireless sniffer? The number of channel are more and the amount of data you're going to acquire is mush higher respect the standard.

- Plain data
- Multimedia contents (video, audio from VoIP services)

You have different types of condition depending on the channel: it can be moving or can be static. That fact that the wireless channel is **time-variant (or time-varying) and frequency-selective** due to **speed condition**, **mobility condition**, and **electromagnetic condition** make more complex the design and analysis of wireless communication systems.

Operations:

- Sense the wireless channels -multiple freq. (many Aps or roaming)
- Consider frames types, sizes, approximate number of frames and bandwidth requirements, motion conditions
- must be able to accommodate the theoretical maximum throughput associated to all the channels (storage, processing power)

FEATURES FOR GOOD WF TOOLS

- 15 radio components (or cards); support channel hop
- use of a GPS (Geographical Positioning System) to get accurate timestamps and outdoor locations (log files)
- capture all traffic without applying any capture filter - filters could be used to collect traffic from specific AP or clients based on their MAC addresses (BSSID or station address).
- completely passive device (hardware attenuator to reduce power: monitor mode for wireless card)
- external antenna connector (to increase the gain)
- remote access
- Side information regarding the data collection (received signal strength or RSSI, capture device, channel, and other signal/noise quality information)

These are (basically) the evidences that are you processing in network forensics.

In ANALYZING EVIDENCES Typically we look for: AUTHENTICATION, ATTRIBUTION, MONITORING.

EXAMPLE: INTRUSION DETECTION SYSTEMS (IDS)

I want to know if there is some activity not authorize in the network, done analyzing logs or network traffic. If I see a strange streams of package respect the normal. IDS can be divided into:

- **Network Intrusion Detection System (NIDS)** monitor at network layer by evaluating network traffic, analyze packet streams (basically sniffers).
- **Host-based Intrusion Detection System (HIDS)** operates directly on the computer hosts by monitoring system logs, system processes, files, or network interface.
- **Protocol-based Intrusion Detection System (PIDS)** monitors a specific protocol (e.g. HTTP).
- **Application Protocol-based Intrusion Detection System (APIDS)** focuses on some specific applications (e.g. monitoring SQL protocols).

How can you do this? Firewalls, IDSs, computer workstations, anti-virus software, databases, end-user applications. The idea is have to an ecosystem of this tools for working together.

Usually they do not work jointly!

HOST-BASED IDS

Monitor

- system activities,
- log files,
- incoming network traffic

- system calls
- audit logs
- error messages

} OS

NIDS detect attack, HIDS tell if they were successful

HIDS detect attacks performed locally (using USB pens, out-of-band attacks, terminals, etc.). You can verify/detect attacks that verify throw different channel not only throw the network.

Out-of-band is a term used to describe any traffic or access that doesn't traverse the public or monitored network (private network, a console connection, or an RS232 serial connection). This makes HIDS robust to Out-of-band attack, normally NIDS or sniffer are not. HIDS are robust to fragmentation and TTL (Time-To-Live) manipulation: *stealth attacks*

NIDS can be fooled by splitting packets into multiple chunks (which are reassembled at the receiving host) or manipulating the TTL.

Problems:

- *Manageability*: Maintenance and configuration are time consuming and inefficient (must be done on each device)
- *Micro view of network attacks*: can't detect common reconnaissance attacks against the host or a range of hosts (ping sweeps or port scans)
- *Compromised hosts*: network connection of IDS can be disabled; no notification is sent to the manager
- *Operating systems' limitations*: hosts may have different OS; IDS software must work on each of them.

NIDS

use probes or sensors installed throughout the network sniffing the network (promiscuous mode) looking for traffic that matches a defined profile or signature. Once the match is positive, the probe sends an alert and can take action to prevent any additional access. Or it's detected intruders throw the statistics, when the packets statistics deviate from normal behaver send an alarm.

See intrusion from a network perspective

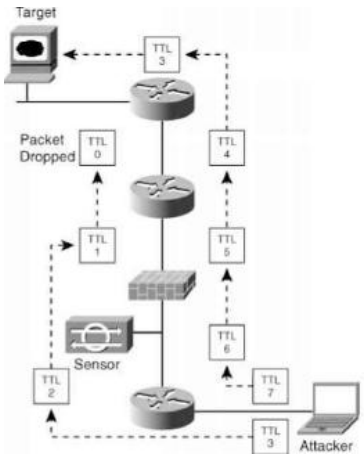
- detect ping sweep or port scan on multiple hosts
- do not depend on the resources of host machines (no waste of network and CPU capabilities; no OS compatibility issues)

Two interfaces: Monitor Interface (MI) and Command & Control Interface (CCI)

Problems:

- ***Bandwidth***: Network probes must receive all network traffic, reassemble that traffic, and analyze the traffic. Network size can increase as well: the NIDS system must scale. This take time.
- ***Packet fragmentation and reassembly***: when your transmitting a packet and this packet it's too large it's fragmented in multiple parts. Hackers conceal their activity from NIDS fragmenting their packets: the packet needs to be reassembled (first to last or last to first) to be analyzed. No problem if no overlapping occurs: hackers send overlapped packets, so they altered the statistics.
- ***TTL manipulation*** (it also altered the statistics)
  - TTL field of TCP/IP packets specifies how long a packet should be considered valid (1-255).
  - Each time a packet passes through a router, TTL ← TTL -1
  - TTL = 0, the packet is discarded.
  - Attackers sends packet with low TTL to create fake traffic hiding attacking packets (high TTL) (in other case it will detected by the firewall)

  N.B. this attack requires a detailed view of the network. It make the statistic normal so no attack will be detect.



- ***Encryption***: Data can not be seen if encrypted: place sensors outside encrypted channel.

HYBRID IDS

Combine the benefits of each type of IDS. Both the sensors and the hosts report to a centralized management or director platform. Need to manage a lot of information. They can incorporate different triggering mechanisms.

IDS TRIGGER

IDS have two triggering mechanisms. IDS are merely packet sniffers: need to teach the IDS what is an attack. You analyze the computer activity when is in normal behavior (no attack).

Anomaly Detection (Profile based)

analyzes computer activity and network traffic looking for anomalies.

Anomaly: deviation from what is defined normal (by system administrator) -> *user group profile*
e.g. you won't expect telnet towards a SMTP server

How to build a user group profile
statistical sampling, rule-based, and neural networks.

Pros:
- Intruders do not know if they generate alarms (attack can not be configured on signatures)
- Detect internal attacks
- Profiles are dynamic

Cons:
- High initial prep time
- No protection during training
- Constant update as users' habits change
- Defining normal behavior can be difficult
- False positives, false negatives
- Hard to understand

Misuse Detection (Signature based)

Use signature files to detect intrusive activity (analogous to signature files commonly used in virus-scanning)

Pros:
- Files created on known attack (if matching, the attack is sure)
- Given the files, protection is instantaneous
- Easier to understand and configure (manager knows what is going on)

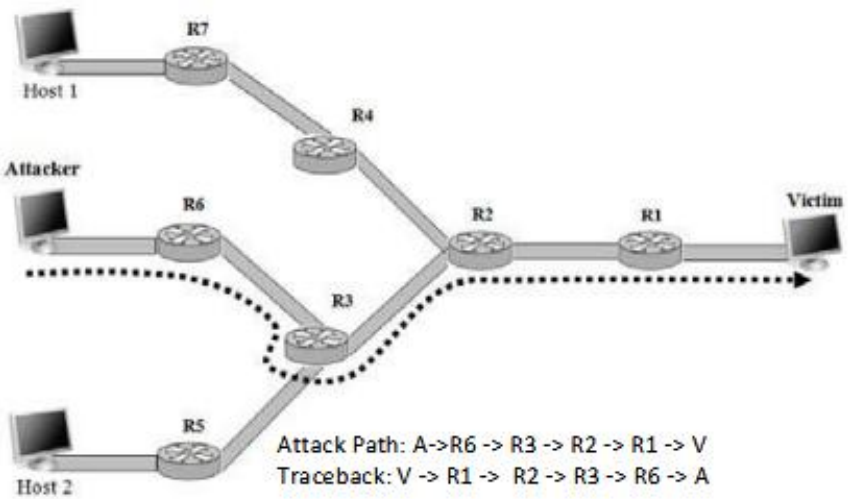
Cons:
- Inability to detect new or unknown attacks
- Inability to detect variations of known attacks
- Signature database administration
- Sensors must maintain state information

TRACEBACK STRATEGIES

Network forensics deals with capture, recording, analysis and investigation of network traffic to traceback the attackers.

Problem: given a packet or a file, find the originating terminal. Used in DDOS attack or against IP spoofing

Characteristics: Time of analysis (real time/post mortem), Data source (flow based/packet based)



Assumptions:

- attackers are able to generate and send any packet
- multiple attackers may act coordinated
- attackers are aware of the traceback ability
- routers have limited processing and storage capabilities
- routers are rarely compromised and all routers may not participate in traceback
- suspicious packet stream are made of a few packets

Requirements:

- compatibility with existing network protocols, routers and infrastructure
- simple and minimal number of functions on transit routers
- support for incremental implementation, partial deployment and scalability
- minimal overhead of time and resources (processing, bandwidth, memory)
- fast convergence involving a few packets
- minimal involvement of the internet service provider (ISP)

TRACEBACK SOLUTIONS

Logging packets at key routers and later mining them for attack path reconstruction:

You have key routers that store packet information, I can correlate strange packet activity that are in different routers for constructing a path.

source path isolation engine (SPIE): hash of multiple fields in packet autonomous management network (AMN): monitoring manager receives requests from sensors

Packet-marking places part or the complete address of the router into the IP packet randomly with a fixed probability or only once deterministically. Some router place some marks on the packets so that you know that packet went throw that specific router.

probabilistic packet marking (PPM)
advanced and authenticated packet marking (AAPM): 8-bit hash of address
algebraic packet marking (APM): algebraic techniques to calculate 15-bit marks as points on polynomials.
fast internet traceback (FIT): uses a fragment of the hash, the number of fragment and a distance field
deterministic packet marking (DPM)
deterministic edge router marking (DERM)

Hybrid traceback approaches integrate packet marking and packet logging to achieve the advantages of both the techniques.

distributed link list traceback (DLLT): router marks a packet, stores current IP address and packet ID in the marking table
hierarchical IP traceback system (HITS)
hybrid single packet IP traceback (HIT)
logging and deterministic packet marking (LDPM)

Autonomous System can be a group of networks regulated by one or more entity, which enforces a clearly defined routing policy.

Traceback strategies are defined here as well.

ATTACK RECONSTRUCTION

In order to refine defenses, it is necessary to analyze and understand how attackers works.
Log files are extremely important!! You need analyze them.

Honeypot refers to a set of services, an entire operating system or even an entire network that is built to lure and contain intruders. A honeypot is a decoy computer system or network designed to lure, detect, and study cyber attackers by mimicking legitimate, vulnerable targets.

- Designed to be (easily) compromised
- Well monitored
- No essential activity is carried on
- A way to draw attack away from more crucial elements in a network
- Useful tool to study attackers.

Problems: From a legal point of view, honeypots can be problematic.

- Honeypot has no value, and therefore, it is not possible to legally claim any damages.
- borderline between keeping attackers out of a network and inviting them.
- may be challenged as an unfair entrapment.

Slide 37-38

LECTURE 6

DEEP WEB APPLICATIONS

Communication that are explicitly hidden

- Virtual Private Network (VPN) with TOR
- Free Anonymous Internet (FAI): used in blockchain because it is a decentralized deep web
- FreeNet: Network of peers to peers and designed to securely and privately spread information on the Internet.
- ZeroNet: torrent with Bitcoin encryption.

TOR NETWORK

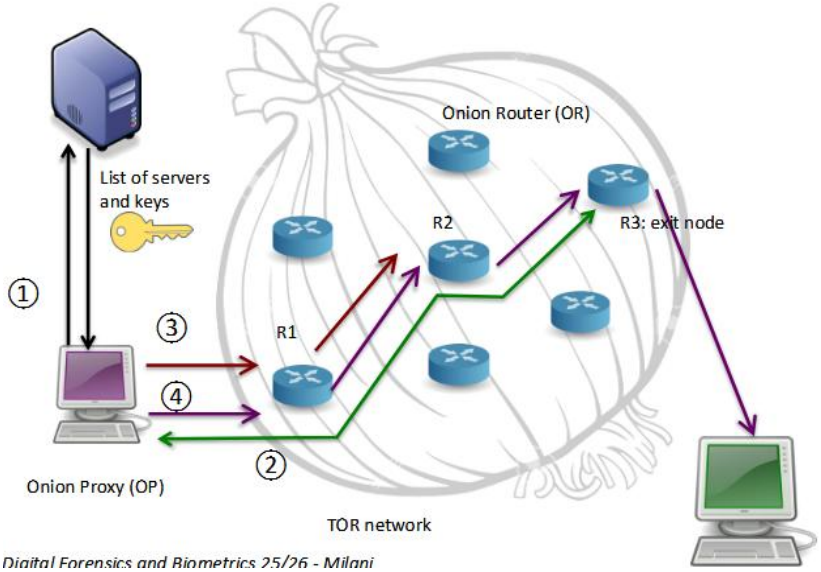
TOR network is a system that allows for anonymous communication online by routing internet traffic through a series of encrypted relays. Protect privacy and anonymity.

• Data are encrypted multiple times (nested encryption)

• Every router correspond to an encryption layer

• No router knows both the destination and the sending nodes.

You have a list of servers provided for users and you have a session key. The idea is that every packet it's encrypted multiple times. Where is the weak point of this network? The exit. There is another weakness I can do correlation with statistic between sender and receiver I can understand "this guy talk with this other guy", I can mitigate this effect using padding (you're adding 0 at the end of the packet, you don't use this bits but you change the length of the packets – but this reduce your throughput).



① Ask for the list of servers
② Verify identity of servers
③ Generate session keys for the session to be initiated
④ Create a connection and transmit Data

Once connection is established, data are exchanged in cells of 512 bytes (to avoid profiling)

TOR FORENSICS

Connection has been created: all messages will be sent as K1(K2(K3(DATA)))

So there are multiple encryption, this means that if you have the key K1 you can't decrypt all the message. This keys are exchanged router by router.

Advantages:

- Everything is encrypted
- No-one knows both sending and destination nodes
- Packet lengths are disguised
- Traffic shaping is possible (sends PADDING)

Weaknesses:

- Timing analysis
- Weakness in the exit node

Strategy	solution	Aims
RAM analysis	Belkasoft RAM capturer Hex dump	types of documents, visited websites, downloaded content
Registry analysis	Regshot	TOR installation, last access, date information
Network analysis	Wireshark, miner information	Monitor network traffic
Database analysis	Locate db files	Gather connection information

OTHER SOLUTIONS

- **Spidering**
  - dark web crawler that collects information on web sites (focused on a specific topic, hidden: automatically filling hidden forms)
  - Features: URL, BoW characterizing documents,
- **Link and content analysis**
  - Measure similarity between contents
  - Methodologies: information sources, domain spidering, backlink search, filtering, indexing and analysis, clustering, activity modelling
- **Dark Network Analysis**
  - Network analysis: topology
  - Graph modelling, small-world property, scale-free property
- **Interactional coherence analysis**
  - Check communication and verify the characteristics of messages and interactions: interaction patterns
  - Lexical relations, direct address, co-reference, interaction patterns from messages
- **Authorship analysis**
  - Writing style features
  - Image/multimedia data analysis
- **Sentiment analysis**
  - Semantic classification of text/images/videos/audio
- **Affect analysis**

THE PROCESSING HISTORY OF A MULTIMEDIA CONTENT

- **Flow of content:** Real scene -> Acquiring device -> Storage -> Post-acquisition editing -> Upload (Social Media like Snapchat, Picasa, Flickr) -> Download -> editing from another user -> New upload -> Download of the new version.

PROBLEMS

For a given image/video/audio file:

- Who took it? Who created it?
- What is its origin?
    - Was it modified? If so, which changes were operated?
    - How much of the original content was altered?
    - Can I trust its content?
    - Is it ai generate or ai edited?
- **Additional questions:**
    - What does this multimedia content tell me about the scene where it was acquired?
    - Is it possible to relate it to other contents?
    - Was the software/device that generated/took it legal?

AUTHENTICATION OF DATA

Given a multimedia evidence, was it acquired by that device? At that time?

- Validation of evidences.
- Tracking of the origin of illegal data:
    - Child pornography.
    - Material promoting illegal activities (terrorism, ...).
- Who owns the data? (Watermarking – visible/invisible).
- Was it manipulated? If so, in what parts? How much of the original content has been changed?

DATA ENHANCEMENT

Tools for image/video enhancement in multimedia forensics.

Images/videos/audios are usually taken in adversarial conditions (e.g., video surveillance at night, eavesdropping in a noisy environment). So usually the quality is low.

- Need to process evidences so that quality is enhanced.
- Perform this in a forensically sound way.
- Localize and isolate events/details from data.
- Identify object/persons/time/locations.

FAKE DETECTION

- Fake photos/videos/audios have been generated since the beginning of photography/music.
- **Reasons:**
    - Artistic reason.
    - Frauds.
    - Manipulate public opinion.
    - Promote trends, aesthetic models.
- Need to detect manipulated data from real ones!

TRACE THE ORIGIN OF A GIVEN CONTENT

- Often content goes viral.
- Multiple copies exist. (for example if you post an image on insta and than downloaded, the downloaded one it's not the original anymore)
- Traceback the origin and the history of a given content.
- *Example categories listed:* Original, Crop, Text Overlay, Watermark, Face Swap, Joystick, Splicing.

OBSERVATIONS

- Lifetime of digital objects is unbounded.
- Objects go through multiple processing steps.
- Processing steps leave characteristic footprints.
- Typically, footprints are deemed as "misprints".
- **MF approach:** footprint as an asset.

Trace back the history of multimedia objects.

HOW MULTIMEDIA FORENSICS TRIES TO ANSWER

- Multimedia contents can be analyzed under different aspects.
- Most of the multimedia files are endowed with **meta-data** related to acquisition device, editing software, etc. (e.g., Exchangeable Image File Format - EXIF tags).
- Forensic analysts can use them (Focal length/exposure, GPS location, Sampling, Editing software like Photoshop). All the social media deleted the most of metadata and replace them with theirs.
- **Problem:** They can be easily removed. An hacker can replace them.
- **Solution:** Look for traces on the pixels of image/videos (or on samples for audio track).

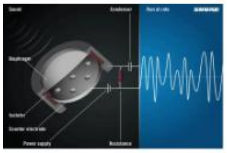
BASICS OF SIGNAL PROCESSING

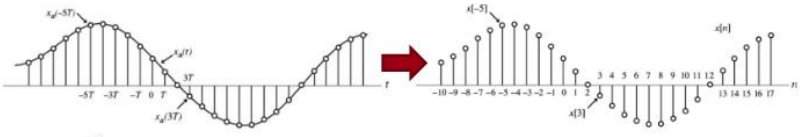
DIGITAL SIGNAL PROCESSING

- After acquisition, digital signals can be modeled from time series or matrices of integer values.

A real-world continuous signals (analog) are transformed into discrete numbers (digital) through the fundamental steps of **Sampling** (breaking the signal into points in time/space) and **Quantization** (assigning a numerical value to each point).

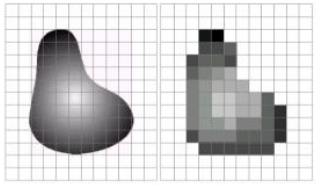
Audio example

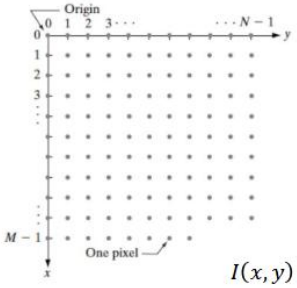




$F_T = \frac{1}{T}$ Sampling frequency $x[n] = x_a(t)_{t\ =\ nT} = x(nT)$

Image example





$$I(x,y) = \begin{bmatrix} I(0,0) & \cdots & I(0,N-1) \\ \vdots & \ddots & \vdots \\ I(M-1,0) & \cdots & I(M-1,N-1) \end{bmatrix}$$

QUANTIZATION

On images for each pixel I have an integer.

- Each sample is quantized and represented with a finite number of bits (8, 12, or 16 bits). (12 and 16 bits are typical are used in medical application).
- Most images use 8 bits/sample: each sample is an integer from 0 to 255 depending on the intensity of the light.
- For color images, digitization is performed on the three components R, G, B.
- Binary signals are present as well (Examples: 254 = 11111110, 150 = 10010110).

BINARY SIGNALS

Some images use 1 bit representation (0 or 1). I find them in biometrics.

- Sometimes whenever shape and geometric details are involved, binary images are used.
- Usually they can be obtained thresholding a standard image or applying some segmentation.
- Examples: fingerprints, silhouettes for gait analysis.
- The signal is a matrix whose cell can assume binary values {0,1}.

$$I = \begin{bmatrix} I(0,0) & \cdots & I(0,N-1) \\ \vdots & \ddots & \vdots \\ I(M-1,0) & \cdots & I(M-1,N-1) \end{bmatrix} \qquad I(x,y) \in \{0,1\}$$

- **Thresholding:** everything is over the threshold is 1 everything is under it is 0.

$$B(u,v) = \begin{cases} 1 & I(u,v) \geq T \\ 0 & I(u,v) < T \end{cases}$$

where T can be computed in different ways.

- **Otsu's method:** ***it skipped on class?***

- **Segmentation:** Segmentation of foreground objects.
  - **Simple case:** define background and Foreground with respect to colors.

More complex solutions: gradient based, watershed, using neural networks.

Genomic Signal Processing

- For DNA/RNA sequences.
- Signal can be seen as a function. $\boldsymbol{s}[n]: \mathbb{R} \mapsto \{A, G, C, T\}$
- Or a set of 4 binary functions (Voss representation) (used for represent a four dimension in binary):

$$\hat{s}_1[n] = \begin{cases} 1 & \text{if } s[n] = A \\ 0 & \text{otherwise} \end{cases}$$

$$\hat{s}_2[n] = \begin{cases} 1 & \text{if } s[n] = G \\ 0 & \text{otherwise} \end{cases}$$

$$\hat{s}_3[n] = \begin{cases} 1 & \text{if } s[n] = C \\ 0 & \text{otherwise} \end{cases}$$

$$\hat{s}_4[n] = \begin{cases} 1 & \text{if } s[n] = T \\ 0 & \text{otherwise} \end{cases}$$

Other Bit Depths

- **Typing biometrics:** user's typing patterns (speed, duration of a single keypress, and how long it takes between releasing a key to pressing the next).
- Signals taking values in the keyboard keys.

DEPTH AND 3D SIGNALS

- 3D sensors are employed as well.
- The acquired signal is usually a depth map D=[d(x, y)].
- Depth value is an integer parameterizing the distance of the pixel from the camera itself. In a depth camera for each pixel doesn't store the light value but collect this instead.
- From depth representation, it is possible to infer a 3D modelling of faces or body.
- *Definition of mesh ???*
- A mesh model needs to be specified by: Vertex, Edges, Faces, Normals.
- Types of meshes: Generic (rare), Triangular (more used), Rectangular.

3D TIME SERIES

- Whenever the identification involves the time modelling of a set of 2D/3D coordinates, n-dimensional time series can be involved. $\boldsymbol{s}[n]: \mathbb{R} \mapsto \mathbb{R}^n$
- **Example:** joints position in skeleton for gait modelling.
- A skeleton is a framework of rigid body "bones" connected by articulated joints.
- Used as an (invisible?) armature for the articulated bodies
- Joint has 0-6 degrees of freedom (DoF): three coordinates and 3 rotation angles.

Given images there are key elements for identifying a person. (for example the characteristic on how I walk can use for identifying). My signal for each of this point is represent in x,y,z and multiply for the number of joins. So is a multidimensional signal.

Skip slide 22

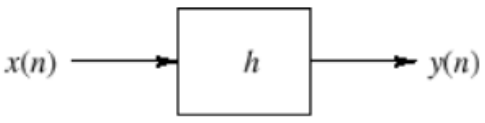
Joints from Acquired Signals

- Usually time series of 2D coordinates (n=2).
- One time series for each joint; traditionally, 15 markers are used for a 30-dimensional time series (Nose, Eyes, Ears, Wrists, Elbows, Shoulders, Hips, Knees, Ankles).
- Acquired via different means:
  - Gyroscopic sensors, accelerometers (IMU).
  - RGB cameras.
  - Depth cameras.

ANALYSIS OF SIGNALS: FILTERING

- Filtering is widely used on biometric signals.

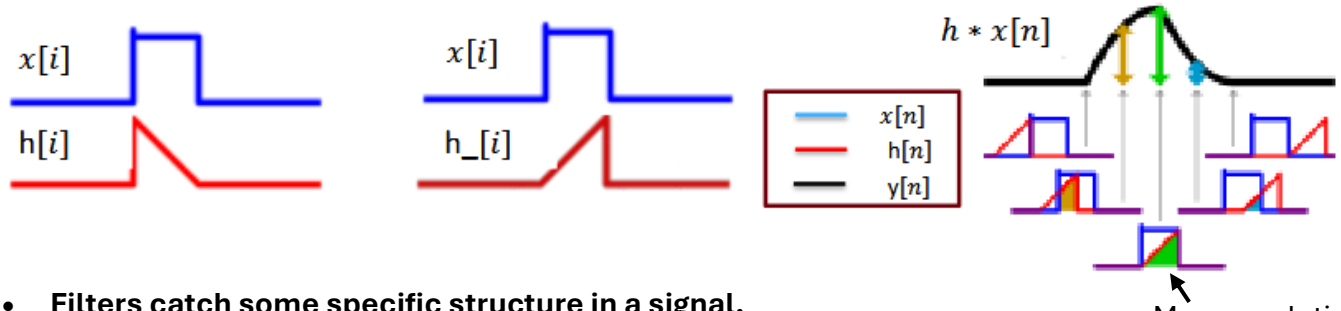
$$y[n] = \sum_{i=0}^{N} h[n-i]\, x[i] = h * x[n]$$



- where h[n] is the response of the filter to the impulsive input $\delta[n] = \begin{cases} 1 & n = 0 \\ 0 & n \neq 0 \end{cases}$

Filtering is the process of removing unwanted components from a signal, such as noise, while retaining or enhancing the desired information. You have two signal: the input (x) and the filter (h called impulsive responsive).
Filtering operation is like an operation when you correlate your original signal (input) with your impulsive responsive signal.

- Ideally, it is like computing the correlation between x[i] and h_[i] with delay n.

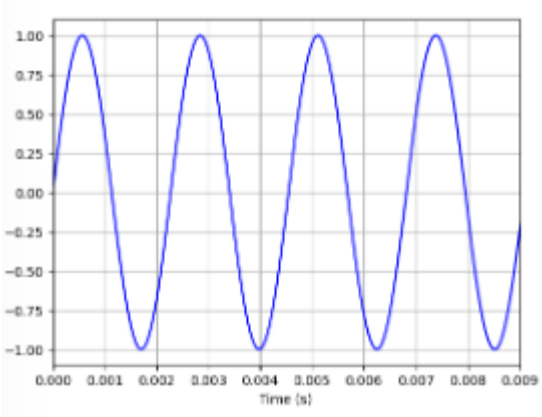


- **Filters catch some specific structure in a signal.**

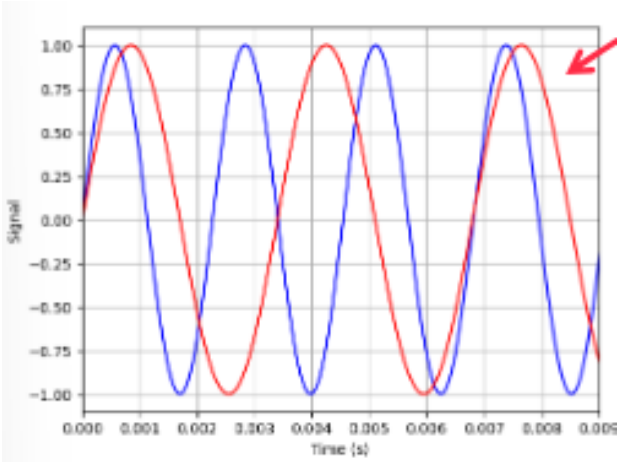
This filtering is very used in binary signal for biometrics in order to have clearer images.
Second operation that we do in **signal processing** is the composition of the signal.
Typically, the signals that represent any signal in signal domain are basically sinusoids.

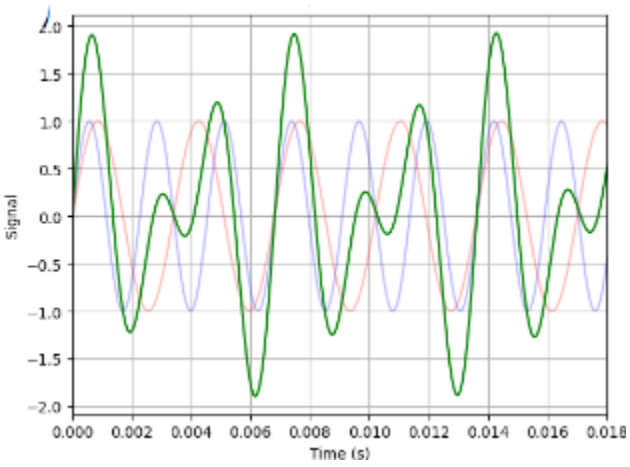
EXAMPLE: 1D SIGNALS (AUDIO)

- Example from sounds but applicable to typing patterns, gait, etc.
- Simple signal: $x_0 = \sin(2\pi f_0 t)$ with $f_0 = 440$ Hz.



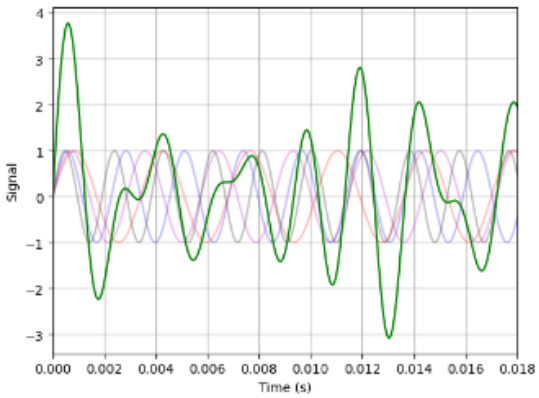
- Adding another note makes it more complex $x_0 + x_1$, $x_1 = \sin(2\pi f_1 t)$ with $f_1 = 294$ Hz.

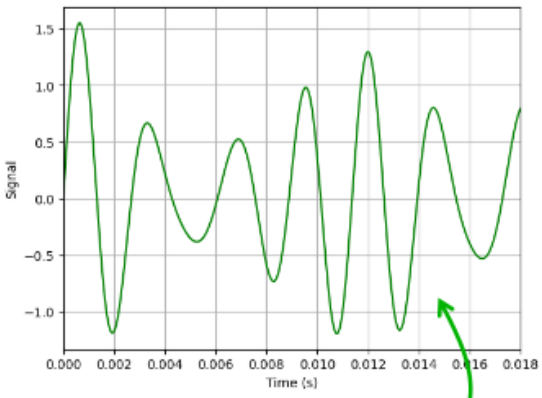




We can do that with “infinite” number of signal and generate as many signal you like.

- Itrìerate multiple frequencies $x_0 + x_1 + x_2 + x_3$
  with $x_i = \sin(2\,\pi f_i t)$ with $f_2 = 349\ Hz$ and $f_2 = 523\ Hz$





- Any signal can be decomposed as a combination of simple sinusoidal components (with different amplitudes

$$0.5\ x_0 + 0.2\ x_1 + 0.8\ x_2 + 0.4\ x_3$$

How to compute the right combination given a signal? -> **Fourier Transform**.
The Fourier Transform can see as a filter, that give you the weight that you need to use in order to create signal out of a sum of weighted sinusoids (or complex exponentials) at different frequencies.

FOURIER TRANSFORM

- Fourier transform can be seen as a filtering with sine-cosine signals.
- Filter with complex functions $\cos(2\,\pi f_0 t) + j\sin(2\,\pi f_0 t) = e^{-j2\,\pi f_0 t}$ and change the frequency f_0 in order to uncover all the weights (coefficients)
- You can use complex signal.
- Change the frequency in order to uncover all the weights (coefficients).

☐ Continuous-Time Fourier Transform (CTFT)
referred as Fourier spectrum or simply spectrum

$$\omega = 2\,\pi f$$

$$X_a(f) = \int_{-\infty}^{\infty} x_a(t) e^{-j\omega\, t} dt = \int_{-\infty}^{\infty} x_a(t) e^{-j2\pi f\, t} dt$$

inverse CTFT $\quad x_a(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} X_a(f) e^{j2\pi\, tf} df$

☐ Discrete Fourier Transform (DFT). (assuming *x* is defined by an array of N values)

$$X(k) = \sum_{n=0}^{N-1} x[n] e^{-j2\,\pi \frac{kn}{N}} = |X(k)| e^{j\theta(k)}$$

Magnitude (dB) $\quad |X_a(f)|$

Phase (rad) $\quad \theta(f) = \arg\{X_a(f)\}$

Inverse Discrete Fourier Transform (IDFT)

$$x[n] = \frac{1}{N} \sum_{k=0}^{N-1} X[k] e^{j2\,\pi \frac{kn}{N}}$$

The Fourier Transform is a representation in frequency where you represent a signal as a linear combination of simple signal/frequency. The representation of Fourier Transform typically follows magnitude and phase representation. In continuous time/domain we are sampling for both the time and the frequency. (that say that in continuous time you can use any frequency, in discreate time no).

Skip slide 29

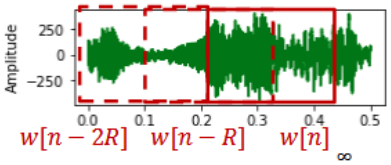
Some examples of continuous Fourier Transform are in slide 30.
Human voice doesn't require high frequency. The higher the frequency I need to use the higher must be the amply factor.

SPECTROGRAMS

If we take the coefficient and plot it as pixel on a matrix, the intensity of the pixel are the absolute value of the coefficient. We see that in a representation called spectrogram.

- Signals can be represented by spectrograms using Short-time Fourier Transform (STFT). You want analyze the frequency as it change throw time.

$$X_d(f) = \sum_{n=-\infty}^{\infty} x[n]w[n - dR]e^{-j2\,\pi fnT}$$



If Constant-Overlap-Add (COLA) propetry is verified $\sum_{d=-\infty}^{\infty} w[n - dR] = 1. \forall n.$ we have $\sum_{d=-\infty}^{\infty} X_d(f) = X(f)$

- It is possible to represent it as an image where rows=frequencies, columns=different windows d (time), pixel intensities =| Xd(f)|
- You will have an idea on how the signals evolve in time (the frequency evolve).

DFT for Images (2D)

The only difference in the images is that you have to add another dimension (2-dimentional).

Continuous

$$I_a(f_x, f_y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} i_a(x,y)\, e^{-j2\pi(\mathrm{f}_x x + \mathrm{f}_y y)}dxdy$$

with its inverse

$$i_a(x,y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} I_a(f_x, f_y)\, e^{j2\pi(\mathrm{f}_x x + \mathrm{f}_y y)}df_x df_y$$

Discrete

$$I(u,v) = \sum_{x=0}^{N-1}\sum_{y=0}^{M-1} i(x,y)e^{-j\frac{2\pi}{NM}(ux+vy)}$$

with its inverse

$$i(x,y) = \frac{1}{NM}\sum_{x=0}^{N-1}\sum_{y=0}^{M-1} I(u,v)e^{j\frac{2\pi}{NM}(ux+vy)}$$

Direct 2D: The term "**direct 2D**" in the context of signal processing most commonly refers to the **direct method for calculating a 2D convolution**. In two-dimensional (2D) signal processing, which is primarily

used for **image processing** (as images are 2D signals), **convolution** is the fundamental operation for filtering, blurring, sharpening, and edge detection. A 2D convolution operation combines two 2D signals:

1. The **Input Signal**
2. The **Kernel** or **Filter**

If you compute the Fourier Transform the most of the energy is in low frequency. In Images the most of the information are in the phase, in the audio the most of the information are in the magnitude.

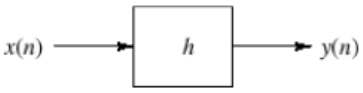
FILTERING

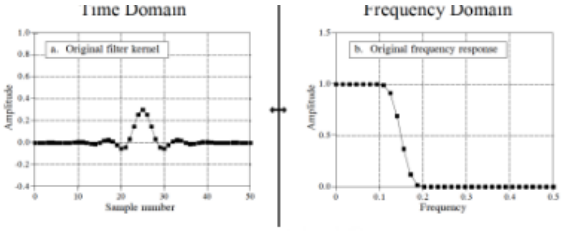
Filtering is an inverse correlation (convolution operation). Allow you to find out where in a signal those parts that looks like as specific signal (the input responder of the filter).

FILTERING IN FREQUENCY DOMAIN

Filtering in time = samplewise multiplication in frequency. Typically you represent filter in Fourier Domain.

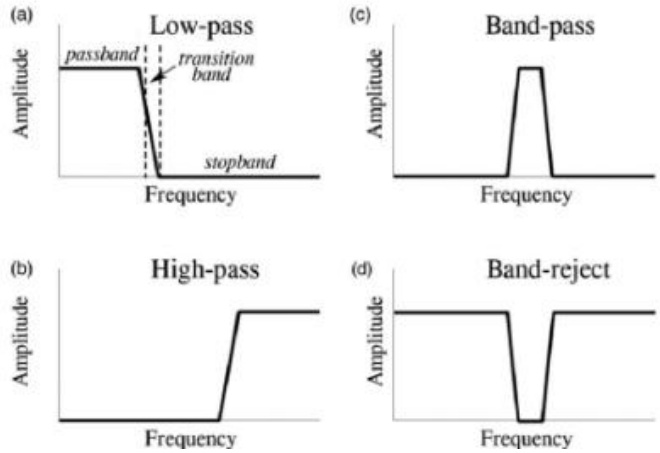
$$H(\omega) = DFT(h)$$
$$X(\omega) = DFT(x)$$
$$Y(\omega) = H(\omega)X(\omega)$$





Usually amplitude response $H(\omega) \in \mathbb{C}$ and can be represented in abs $|H(\omega)|$ and phase $\angle H(\omega)$

Types of filters: Low-pass, Band-pass, High-pass, Band-reject. (the 2 last are called also Narrow filter)

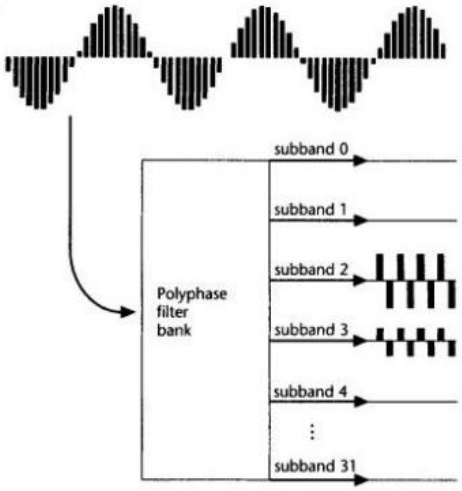


You can remove frequency for better audio: for example if I have an audio with a sound back ground of a conditioner I can remove the frequency of the conditioner to have a better result.

Examples of Filtering

- Filter responses are maximum whenever input signal and response are aligned and very similar.
- Example: Average filters (smoothing).
- **Filtering for Biometrics:** (high light some specific part)
    - Helpful in ridge extraction for fingerprints (Gabor filters).
    - Used to fix broken ridges, separate joined ridges, or enhance pores.
    - Also used for audio denoising.

SUBBAND DIVISION



It is possible to use filtering to characterize a signal.

- Subband coding encodes a signal/image by breaking it down into the sum of several simpler components. (mp3 audio does that)
- **Polyphase filter bank:** The original signal is filtered and each frequency subband is isolated.
- **Images:** Each component related to details.
- **Sound:** Each component related to different objects/people.
- If frequency ranges are small enough, the accuracy is high.

Suppose you have the Fourier behavior of your signal and you divide it throws multiple rectangular filters, you can add more and more filter until you get a good approximation of the signal itself. The more filter I use the better the representation will be but the problem is the complexity is higher.

LECTURE 8

DIGITAL CAMERA MODEL

Light it object, the object reflect light, the lens focus this reflection on a matrix sensor, then the sensors convert this light intensity in a voltage (number), this number is process and save in a memory (digital one). All this steps here leave some traces, they can be used to trace back to the origin of the image.

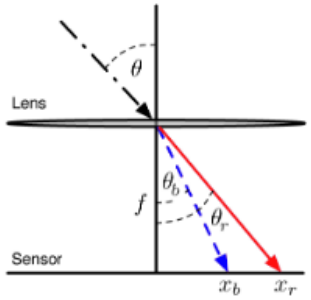
- Projective model is applied from object points to analog optics.
- Image plane (real) is associated to a matrix of finite elements (pixel): each element is a digital sensor (CCD o CMOS) that converts light intensity into a number (ADC).
- Matrix of photoresponsive diodes which turns photons into electrons.

LENS ABERRATION

- Most images contain a variety of aberrations that result from imperfections and artifacts of the optical imaging system. This because lens goes to a fabrication model that is not perfect.
- You can have distortion and color aberration. Can be relate to the material and the geometry of the lens.

Fresnel's equation. $n \sin\theta = n_f \sin\theta_f$

θ : incident angle
θ_f: refraction angle
n, n_f: refraction indexes. (depends on wavelength λ)



- When you apply this of any kind of material colors separate.
- Different colors = different λ = different refractions.

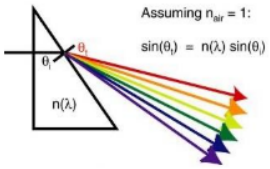
$n \sin\theta = n_B \sin\theta_B = n_R \sin\theta_R$

Since n are different angles must compensate $\theta_R \neq \theta_B$

Dividing by $\cos\theta_R \approx \cos\theta_B$ we have

$$n_B \operatorname{tg}\theta_B = n_B \frac{x_B}{f} \approx n_R \frac{x_R}{f} = n_R \sin\theta_R$$

$$x_R \approx \frac{n_B}{n_R} x_B = \alpha x_B$$



CHROMATIC ABERRATION

This is not uniform to the lens, on real lenses they uses compensation. The **Chromatic Aberration** is more evident on the side of the images, where the curvature of the lens is higher and the light arrive with a worst angle.

- Different wavelengths converge to different points.
- Examples: Near Lens Center vs Near Lens Outer Edge.
- Can be compensated (achromatic doublet).
- **Model:**
  - From previous equations we have $(x_R, y_R) \approx \alpha(x_B, y_B)$

- A more complex model adds a center of aberration:

$$x_R = \alpha(x_B - x_0) + x_0$$
$$y_R = \alpha(y_B - y_0) + y_0$$

  This model tells us that more you get far from the center of the lens the more the colors have different values.

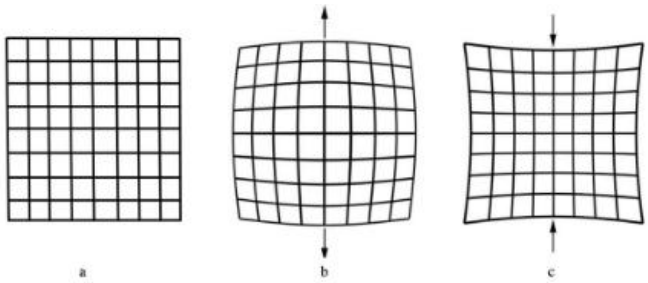
- Using green as reference, find parameters α, x_0, y_0.

RADIAL DISTORTION

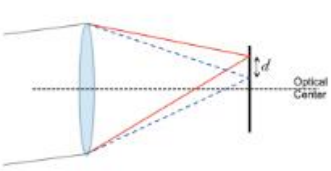
- All the lenses show this (compensated on the firmware).
- Utterly evident for fish-eye lenses (the image is curved).
- For omnidirectional cameras, use together with tangential distortion. The distortion is a problem in case you want to measure or analyze the images.

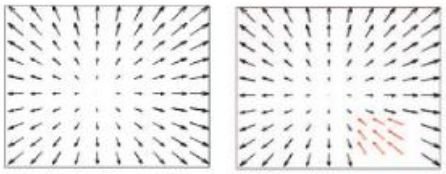
LENS FORENSICS

- For forensics this distortion are useful because you can build a model and defined how much the lines are distorted.
- Most lenses introduce different kinds of lens aberrations such as spherical aberration, field curvature, lens radial distortion and chromatic distortion.



- Aberrations of sphericity are due to the building process: by calibrating the camera it is possible to estimate distortion parameters. For example if you have a image and you estimate k1 and k2 than you can associate which is the camera as acquire that specific image.
  - features for identification
  - Formula: $r_u = r_d + k_1 \cdot r_d^3 + k_2 \cdot r_d^5$
- Similar considerations for colors.





PRINTER AND SCANNER FORENSICS

Printer and scanners have lenses, you have a lots of elements in those. For example in scanner we have to consider the effect of the gears, in printers we to have think about all the chemical issue due to the ink.

Printing Technologies

- **Inkjet Printing:** Ink, Nozzle, Droplet, Substrate.
- **Laser Printers:** Charging, Laser Exposure, Toner, Developing, Fixing, Transfer.
- **Laser Scanning:** Scanning Lens, Objective, Laser, Photoconductor.

Given a specific devices you will have different traces that are left on the image by different factor:

Scanners' signatures:

- Sensor noise:
  - Denoising filters.
  - High-frequency wavelet coefficients.
  - Neighborhood prediction errors.

Laser printers' signatures:

- Gear mechanism.
- Polygon mirror wobble.
- Optical photoconductor (OPC) angular velocity.

Ink-jet printers' signatures:

- Dot shapes. (an id and a time stamp)
- Dot placement.
- Ink-jet nozzle, ink chemistry.
- Periodic variation by missing jets or paper-advance errors.

SENSOR NOISE

You can estimate noise because it is a physical process, it's not control, it something is kind of unique, it's not systematic but in many situation is systematic.

DIGITALIZATION

This type of noise is systematic.

- **Idea:** the energy of incoming photons is converted into a voltage thanks to the photoresponsive material.
- Voltage is proportional to the integral energy absorbed by the sensor (photoelectric effect).

In digital camera there are two different methods to acquire this information: CMOS and CCD.

CMOS ACTIVE-PIXEL SENSORS (APS)

You have a process that's reads the charges information (light) by row and column order, then this charge some resistor that create a voltage that is read in row and column.

1. Every pixel = photodiode + readout amplifier.
2. This allows converting the energy stored as charge in the photodiode in tension and amplify it; the value is then sequentially read according to a row-column ordering.
3. 1 pixel = 3 transistor: conversion electronic charge/tension, reset of photodiode, reading.

4. Sensor array is organized in a checkerboard structure; dedicated circuitry is used to control the sequential reading.
5. Address can be specified with (x,y) coordinates.

CCD (Charge-Coupled Device)

Here instead of acquiring and reading the tension we generate some charges and then these are capture using the shift register technology.

- Two components:
- Photoresponsive region.
- Shift register to transmit the acquired value.
- **Pixel acquisition process:**
    1. Light hits the capacitor of the photoresponsive region.
    2. Capacitor starts charging: the amount of charges depend on the intensity.
    3. When acquisition time ends, the charge is transferred to the nearby capacitor (shift register) using a control circuitry.
    4. The last capacitor is discharged using an amplifier that changes the charge in a voltage. Voltage is sampled and memorized.

The difference is that in CMOS the reading is sequential in this case the reading is attending.

PHOTOSENSORS IN A CAMERA

- Similar to buckets collecting rain (photons) until a certain level (the pixel value) is reached.
- **Ideally:** When uniform light falls on a camera sensor, each pixel should output exactly the same value.
- **Practically:** Small variations in cell size and substrate material result in slightly different output values.
- Some noise is introduced into the image:
    - **FPN:** Fixed Pattern Noise.
    - **PRNU:** Photo Response Non Uniformity.

You expect the value to be uniform because the light is uniform but the value pixel by pixel are different. This is because during the fabrication process there are imperfections. This different variation it's like DNA for cameras.

CAMERA BALLISTIC

- Like a gun on the bullet, the acquisition devices leave distinctive traces on the acquired content.
- **What for?**
    - Distinguish if the content is real or computer generated.
    - Distinguish between different devices (camera, scanner, etc.).
    - Distinguish which camera captured the image: **Camera identification**.
    - In legal content this is useful.

PHOTO RESPONSE NON UNIFORMITY (PRNU)

The ai generated image leaves a sort of PRNU, in this case depends not only on the network but also on the training (dataset and architecture).

I_0 is the ideal image (without noise). At this I add a noise I_0K, this noise is depending on the signal. This is called multiplicative noise. Θ all the other noises.

- PRNU is caused by the different sensitivity of the sensors to light.
- Due to the manufacturing process - does not depend on temperature and time.
- **Model:**

$$I = I_0 + I_0K + \Theta$$

I_0 sensor output without noise

K PRNU fingerprint of camera C (multiplicative noise)

Θ all the others noise terms (shot, readout, etc.)

- **Extended modelling:**

$$I = g^\gamma [I_0(1 + K) + \Theta]^\gamma + Q$$

- where gathers compression/quantization noise and is the gain factor. This is a better model because sometimes the firmware of the camera organizes the noise (for example the light of the camera).For jpeg image this model is better because you use the noise compression.

How does the PRNU Estimation work? Imagine that you have a set of images, taken with a specific camera, you performing a filter (any kind) that's gives you a noise free image. We called this noise free image I^, and I consider a noise version of the image an estimation of I_0. Then I compute I-I^. So if I have multiple images I use the ∑. *Ask more about this equation to Gemini.*

PRNU Estimation

1. Denoise the image $\hat{I}^{(0)} = F(I^{(0)})$.
2. Compute difference $W_I = I - \hat{I}^{(0)} = IK + \Sigma$. (Repeat for many images $I_1, I_2, ..., I_N$ generating $W_1, W_2, ..., W_N$).
3. Estimate K:

$$\hat{K} = \frac{\sum_{i=1}^{N} W_i I_i}{\sum_{i=1}^{N} (I_i)^2}$$ Compute noise

PRNU Detection

- **Recipe for a good estimate of PRNU:**
    - The estimation is done on flat-field images (uniform content, low variance).The Images are making in a control environment.

    - The luminance of the images should be as high as possible but not saturated.
    - About 20 flat-field (bright) images suffice to have a good estimate of K.
- **Verification:**
    - Let be another image from the same camera (or another one).
    - The presence of can be determined by a correlation detector.

You take this image J, compute W_J and then you perform the correlation between W_J and K^J. You can do the correlation by:

- NCC (Normalized Cross-Correlation)
- PCE (Peak-to-Correlation Energy). This one is preferred because is more reliable in metrics and in term of authentication.

$$corr(W_J, \hat{K} \cdot J)$$

NCC (Normalized Cross-Correlation)

$$\rho = \frac{(X - \bar{X}) \cdot (Y - \bar{Y})}{\|X - \bar{X}\| \cdot \|Y - \bar{Y}\|}$$

$$X \cdot Y = \sum_{i,j} X[i,j] \cdot Y[i,j]$$

$$\|X\| = \sqrt{X \cdot X}$$

PCE (Peak-to-Correlation Energy)

$$PCE = \frac{\rho^2}{\frac{1}{MN} \sum_i \sum_j NCC_{i,j}{}^2}$$

$$NCC_{i,j} = \frac{\sum_{k=0}^{m-1} \sum_{l=0}^{n-1} (X(i,j) - \bar{X}) \cdot (Y(k+i, l+j) - \bar{Y})}{\|X - \bar{X}\| \cdot \|Y - \bar{Y}\|}$$

PRNU in Practice (Process Flow)

- Smartphone under analysis -> Take some flat-field images -> Estimate PRNU / Pictures under analysis -> Extract noise
- -> Correlation test -> Examine results Conclusion: match/no-match.

Since the peer review is a noise which depends on the intensity it should be in the middle, not to dark not to bright. If the images are too dark you have the problem that the intensity is very low, and the impact of the other noise sources is much higher. If you have a too bright image you are too close to top scale of your image, saturation effect.

Slide 27-28-29-30

SPECTRUM OF COLORS

- Light appears to be white but it is made of many colors.
- Usually coded into 7 colors, but many frequencies are present.
- Cameras usually acquire 3 colors: R, G, B. You can obtain every color doing the linear combination of this tree.
- Your *CCD* don't care about colors, it computes only the intensity

COLOR SENSORS

- Sensors cannot distinguish colors (i.e., wavelength/photons' energy), they count photons' number.
- CCD/CMOS matrices generate images in grayscale. To get colors, we need filters. Thanks at the filter in front of the *CCD* you can recover the color information. For each pixel of the camera you acquire only one between R, G, B.
- So, how is it possible to record 3 value for each pixel? Thanks to the interpolation (Demosaicing)
- **Two possible technologies:**
    1. **3-chip color:** incoming light is divided into the three components (R,G,B) using filters and prisms. Three different sensors are then used.
    2. **Single-chip color:** uses color filters before the acquisition, which separates the different components for each sensor. The complete information is then reconstructed by firmware (Demosaicing).

Biomask: the red green blue value is called biomask. *(non ne sono sicuro)* Mask can be different

CFA INTERPOLATION DETECTION

Interpolation is basically a strategy that we follow in order to find out the missing value.

- Digital cameras acquire a single frequency component per pixel according to a regular pattern: **Bayer mask**.
- Every vendor uses its own Bayer masks, and the interpolation is performed with proprietary and arbitrary hardware.
- *Ask to Gemini the general idea of interpolation*
- There are 2 way for doing **color interpolation**:
    - Vertical (vertical green, vertical blue)
    - Horizontal (horizontal green)

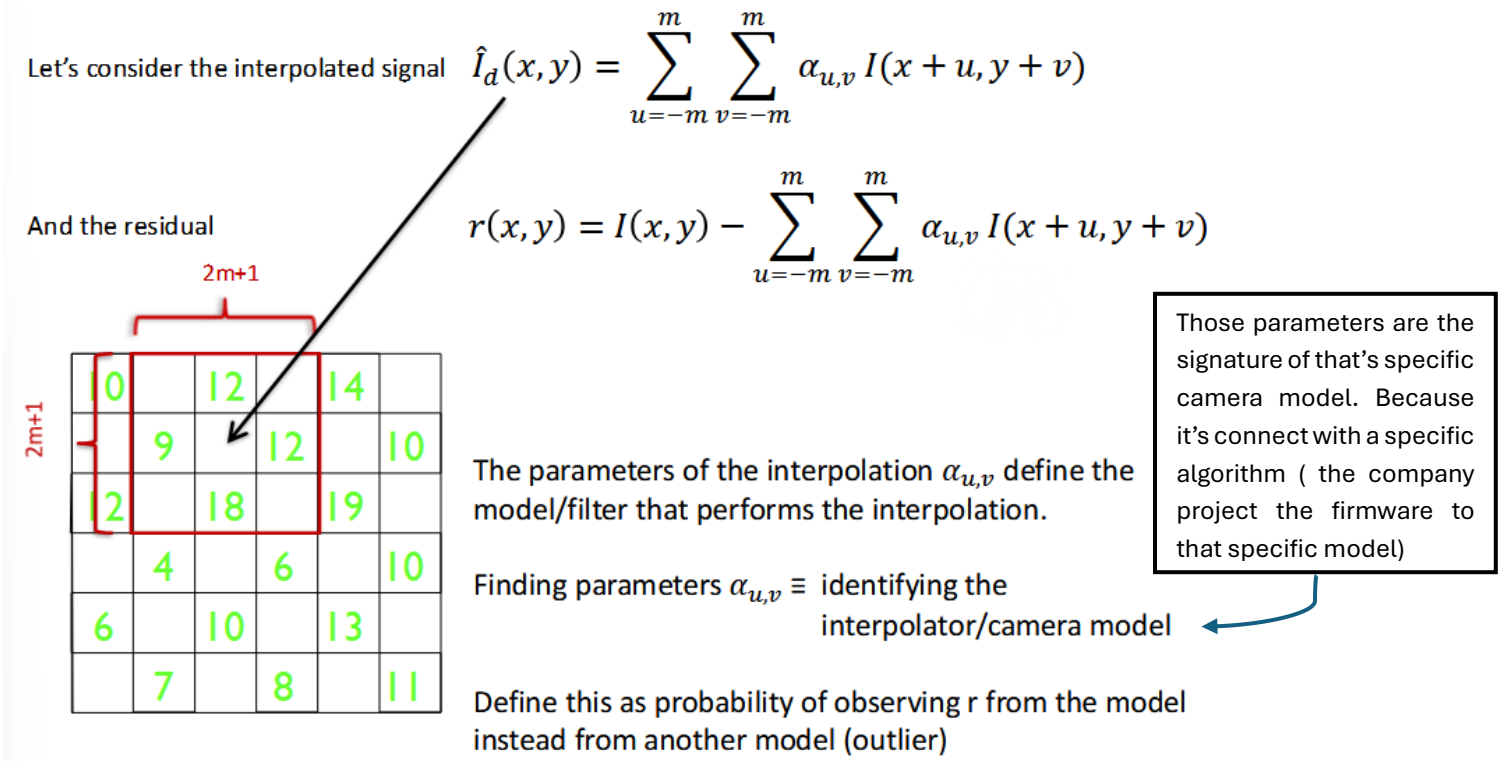
What is the best strategy? Depends, the camera vendor choose. Because the interpolation are decided by company. This is an advantage for Forensics analysist.

DEMOSAICING DETECTION

- We can distinguish images taken by a camera and by a scanner.
- We can understand whether an image or a portion of an image has been resized.
- Interpolation can be inferred from traces left on the image.

A Simple Estimation Approach

- Simple linear model for the periodic correlations introduced by CFA interpolation: weighted sum of pixels in a small neighborhood.
- Both easy to parametrize and can reasonably approximate CFA interpolation algorithms.
- Let's treat colors independently.
- If correlations are known (i.e., the parameters of the linear model), then it would be straightforward.
- **Problem:** There are unknown information: whether pixel is interpolated or not!



You want to find out the parameter $\alpha_{u,v}$ so that this error here (*idk*) is minimized.
You want to find out which interpolation you have to apply to the signal you can approximate this interpolation by a linear combination.

In image we have real pixel (actually acquired on the scene) and interpolated pixel, they come from two difference sources, one the sensor the other the interpolation algorithm. How I know which pixel is interpolated and which no? If you not consider this you could have error. For distinguishing we use EM.

EXPECTATION-MAXIMIZATION (EM)

- Method to find model with missing data (distinguish standard pixels from interpolated ones). Suppose you have a set of points that can be generated by two sources A or B (Gaussians). We can perform this estimation with 2 alternatives steps. Those two are repeated multiple times.
- **Algorithm:**
    1. **E-step** (Estimation step)**:** Given the initial parameters it estimates the probability for each point (distributes values between model/outlier).
    2. **M-step** (Maximization step)**:** Recompute the models (maximize likelihood). So with this new probabilities you fix your probability and you recompute the parameters of the model A and the model B so that this probabilities is a better match.
- So whit this algorithm we solve the problem and we can estimate which pixel is interpolated or not.
- Method to find model with missing data
- We do not know how to distinguish std. pixels or interpolated with the unknown method
- EM solve: introduces unknown or hidden variables
- We introduced a random variable *z*, the value of my pixel is *x*, and the set of parameters is *Θ*.
- Instead of using probability we maximize the log. From the Bayes rules we leads to the result.

Maximize $\mathrm{p}[x;\theta] = \sum_z p[x;z,\theta]$ leading to log-likelihood for samples x_i

$$\ell(\theta) = \sum_i \log p[x_i;\theta] = \sum_i \log \sum_{z_i} p[x_i;z_i,\theta]$$

From Bayes we have

$$log\ p[x;\theta] = log \sum_z p[x;z,\theta] = \log \sum_z q(z) \frac{p[x;z,\theta]}{q(z)} \geq \sum_z q(z) \log \frac{p[x;z,\theta]}{q(z)}$$

That leads to

$$log\ p[x;\theta] \geq \sum_z \sum_i q(z_i) \log \frac{p[x_i;z_i,\theta]}{q(z_i)}$$

$q(z)$ is the posterior distribution; typically $q(z) \sim p[z_i|x_i,\theta]$

$$log\ p[x;\theta] \geq \sum_z \sum_i q(z_i) \log \frac{p[x_i;z_i,\theta]}{q(z_i)}$$

$q(z)$ can be approximated as

$$q(z) \sim p[z\,|x,\theta] = \frac{p[x,z|\theta]}{p[x]} = \frac{p[x,z|\theta]}{\sum_z p[x\,|z]\ p[z]} = \frac{p[x\mid z,\theta]\ p[z]}{\sum_z p[x\,|z]\ p[z]} = \frac{p[x\mid mod]\ p[mod]}{\sum_z p[x\,|mod]\ \ p[mod] + p[x|out]\ p[out]}$$

Being model and outlier equally probable

$q(z) \sim \frac{p[x\mid mod]}{\sum_z p[x|mod] + p[x|out]]} = \frac{L(mod)}{L(mod)+L(out)}$ distributes values between model/outlier

Then, algorihm becomes

1) E-step: compute $\sum_i p[z_i\,|x_i,\theta_i] = \frac{L(mod)}{L(mod)+L(out)}$

2) M-step: maximize $\sum_z \sum_i q(z_i) \log \frac{p[x_i;z_i,\theta]}{q(z_i)}$

Repeat multiple times until convergence

Maximize $\mathrm{p}[x;\theta] = \sum_z p[x;z,\theta]$ leading to log-likelihood for samples x_i

$$\ell(\theta) = \sum_i \log p[x_i;\theta] = \sum_i \log \sum_{z_i} p[x_i;z_i,\theta]$$

From Bayes we have

$$log\ p[x;\theta] = log \sum_z p[x;z,\theta] = \log \sum_z q(z) \frac{p[x;z,\theta]}{q(z)} \geq \sum_z q(z) \log \frac{p[x;z,\theta]}{q(z)}$$

That leads to

$$log\ p[x;\theta] \geq \sum_z \sum_i q(z_i) \log \frac{p[x_i;z_i,\theta]}{q(z_i)}$$

$q(z)$ is the posterior distribution; typically $q(z) \sim p[z_i|x_i,\theta]$

LECTURE 9

COMPRESSION – Coding

Image is saved in compressed form (jpeg), that's means the pixel al shrimp. Only high quality contain images in a non-compressed format. Videos are always compressed. You can traceback to which type of codec it was used. Every camera have his codec, the problem is when the picture is uploaded, because every website have its codec. (for example Netflix, YouTube have their own codec that are different, so the compression on the same media can be different).

- Images and videos are usually compressed.
- Coding standards define bit stream structure and decoding operations.
- Encoder is free; different vendors perform different choices.
- **Codec related features!** (Image, Video, Audio).

LOSSLESS AND LOSSY CODING

Jpeg is a lossy compression, it has some loss.

So you can have small size at the prize of some distortion on the image. Bigger compression lower quality.

- **Lossless:**
  - Data are converted into a binary bit stream.
  - Mapped into non-ambiguous strings.
  - Sources are perfectly reconstructed, no distortion is present.
- **Lossy:**
  - Obtaining a high-fidelity reconstruction (as much as possible) given a maximum limit to the maximum amount of coded data (bit rate).
  - Given a target distortion level between the original source and the reconstructed one, generate a bit rate size as small as possible.
  - Not invertible.
  - High compression ratio.

Example

- QF=90, PSNR=41.26, bpp=2.31, Compression ratio = 1:10.
- QF=10, PSNR=31.98, bpp=0.31, Compression ratio = 1:77.

Lossy Coding Measures and Formulas

- **Objective measures:** you can measure quality using those metrics
  - Error: $e = I(x,y) - IR(x,y)$
  - PSNR: $PSNR = 10log10E[e2]2552\,\frac{255^2}{E[e^2]}$
  - SSIM: $SSIM(x,y) = \frac{(2\,\mu_x\,\mu_y + c_1)\,(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)\,(\sigma_x^2 + \sigma_y^2 + c_2))}$
- **Subjective measures:** Perceptual quality tests.
- **Lossy image formats:** JPEG, Gif, JP2000.
- **Lossy video formats:** MPEG 1, 2, 4; H.26x (or AVC, HEVC); Flv, Ogg, Webm.

Transform Coding for Images – slide 8 skipped

JPEG CODING SCHEME

- **Steps:**
    1. RGB-YUV (color)Conversion + subsampling.
    2. Divide the image into blocks. Block partitioning (8x8 Pixel Blocks).
    3. DCT (Discrete Cosine Transform). (similar to Fourier Transform)
    4. Quantization. Organize the information in bitstream **(this is the lossy part).** Up to this point the process can be inverted, but after quantization you can't because you lose information.
    5. Entropy Coding (Huffman or Arithmetic Coding).
- **Process:** Start of Frame -> Frame Header -> Scan Header -> Segment Header -> Block -> End of Frame.

Frequency Representation

- It is possible to transform an image/audio file in a representation that is more suitable for the compression.
- Frequency-based representation.
- Usually operated on blocks.
- The quantization steps (more or less aggressive) depends on the position of the coefficient
- (Original pixel data vs DCT coefficient data tables).

Quantization & Blocking Artifacts

- Quantization introduces information loss but permits reducing the bit rate!
- **Formula**: $x_q = \lfloor x/q \rfloor$. You take the integer approssimation.
- The quantization of decorrelated data introduces image distortion, called **blocking artifacts**.
- Since coders operate on square blocks, quantization artifacts have a square shape.
- Many coefficients quantized to zero **ringing artifacts**.
- The more aggressive the quantization, the more evident the blocking effect! (the more you reduce the bit rate)
- Bit rate limits impose stronger artifacts if the video/image is difficult to code.

Frame Building

Images is built into a file that is dived in containers.

- **Frame header:** image size (width/height), Components (RGB, YCM,…), Digitization format (4:2:2,…).
- **Scan header:** Identity of the components, #bit/component, Quantization tables per component, Huffman tables (not default).

SIMILAR STRATEGIES FOR AUDIO: MP3

The only difference that we have in audio is that instead of using transformation we use filtering.

- MPEG-1 Audio Layer 3 (MPEG = Moving Picture Experts Group).
- Early '90s.
- Music, audio tracks, recordings. Popular on the Internet.
- **Encoder steps:**
    - PCM Audio Input Time to Frequency Mapping (Filter Bank).
    - Psychoacoustic Model.

    - Bit/Noise Allocation, Quantizer and coding.
    - Bit stream Formatting Encoded Bitstream.
- Frame division, frequency optimization and compression, pseudo-acoustic models.

PREDICTION FOR COMPRESSION

There is correlation between samples, this is not random but there a sort of memory in the signal. We can exploit this memory for our use.

- Acquired video needs to be compressed in order to be transmitted.
- Remove redundancy Prediction is the key!
- Example: 8 bit/sample x 1000 sample = 8kbit.
- If samples are highly correlated, more effective coding can be done, e.g., DPCM.
- Difference between adjacent samples is +1, 0, -1 (only two bits are needed).
- Periodically unpredicted elements are included.

Generic Predictive Video Coding Architectures

- Adjacent frames in a video sequence are correlated: there is a lot of redundant information.
- It is possible to reduce the redundancy predicting the information: only the difference needs to be coded.
- **Process:** Reference Frame + Prediction error = Current Frame. So I don't need to specified the full frame I only need to specified the difference between the prediction and the original one.
- Uses Search Window, Motion Vectors ().

Bit Rate Control

- **Constant Bit Rate (CBR):** Rate control/optimization algorithms are among the digital fingerprints.
- **Constant Quality:** Variable size based on content complexity (e.g., CRF=45 vs CRF=27).
- Depending on the video the bitrate is important. In case a static scene you can use low bit rate, in case a more complex scene you need higher bit rate. More bitrate more memory occupied.

Coding Current Image and Video Compression Standards

- **JPEG:** Continuous-tone still-image compression. Bit Rate: Variable.
- **H.261, MPEG-1, MPEG-2:** Video telephony, storage media (CD-ROM), Digital TV. Bit Rate: 1.5 – 20 *Mb/s*.
- **H.263:** Video telephony over PSTN.
- **MPEG-4, JPEG-2000:** Object-based coding, synthetic content, interactivity.
- **H.264 / MPEG-4 AVC:** Improved video compression. Bit Rate: 10's to 100's *kb/s*.

JPEG Compression Footprints

- Detection of the codec, you can understand if the quality of the image was higher.
- Detection of JPEG coding footprints:
    - **Blocking artifacts** [Liu and Heynderickx, ICASSP 2008].
    - **DCT coefficient statistics** [Fan and de Queiroz, TIP 2003].
    - **Ringing artifacts**.
    - **Image codec detection** (JPEG / wavelet / DPCM) [Lin et al., TIFS 2009].

How did you detect compression strategies? Supposed that you take your tcp coefficient, if the image are not compressed the statistics is normal, if it's compressed the statistic of the coefficient is quantized. The

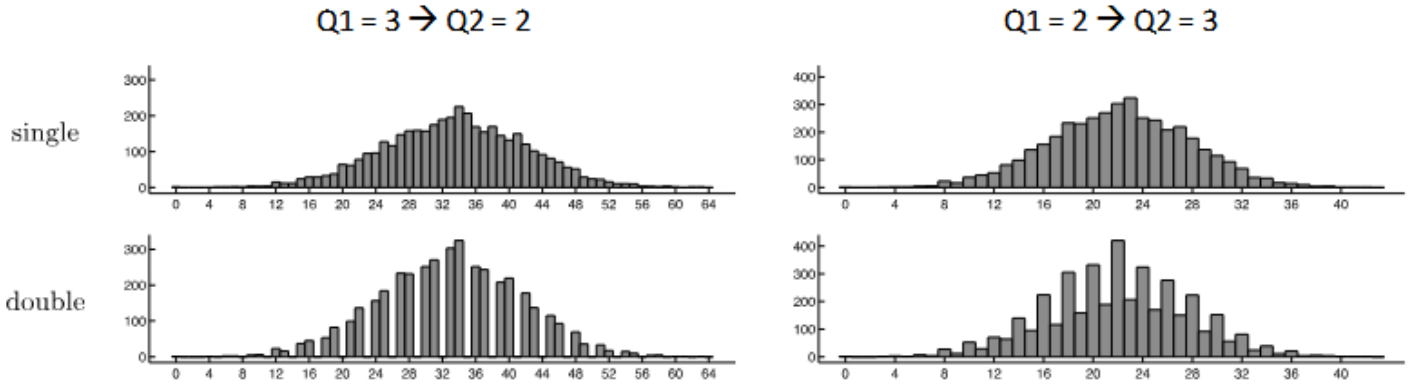
value of the coefficient will assume value within the quantization breed. “They lying on come line statisctic”.

- DCT coefficient statistics
  - Detect if JPEG compressed (or double compressed)
  - Estimate quantization matrix template $p(Y_r;\Delta)=\sum_k w_k\delta(Y_r-k\Delta)$

DOUBLE (MULTIPLE) JPEG COMPRESSION

I perform twice the compression.

- The goal is to identify traces of previous compression and, possibly, to estimate the codecs.
- Detect if single or double compression was applied (and which were the parameters).
- **Methods:**
  - DCT coefficient statistics (Histograms showing gaps or peaks due to double quantization).
  - First digit's statistics.
  - Example: to the left you have this kind of result if the second quantization steps are lower than the first quantization. On the right we have the opposite. So if you detect one of this two on video you know the video was quantizite twice (so is not the original one).



Social Media Statistics

Using the compression datas you can identify the province of the image (with social media)

- Images, during upload on SN (Social Networks), are usually re-compressed.
- Different quantization strategies adopted statistics change.
- Neural networks can distinguish image provenance (e.g., FusionNet, 2019 detecting Facebook vs Twitter vs Flickr).

DIGITAL AUDIO RECORDERS

DIGITAL AUDIO

- Sound is a vibration that propagates as an acoustic wave, through a transmission medium such as a gas, liquid or solid.
- This vibration can be modelled as a wave signal that propagates through space.
- Whenever the wave encounters an acoustic receiver (microphone, human ear, ...) an acoustic membrane starts oscillating.
- In a microphone, oscillation is converted into a voltage.

- In a human ear, vibration is converted into an electric stimulus.
- Voltage signal can be sent to a loudspeaker: Voltage is converted into a vibration of speaker's cone (reverse process).
- Voltage is converted into a vibration of speaker's cone -> vibration back in the air

DIGITAL AUDIO ACQUISITION

- After acquisition, audio (voltage) waves need to be stored in some numerical format.
- **Process:** Sound MIC (converts sound into voltage) -> Voltage amplified and acquired using **ADC unit** -> ADC converts the signal into a sequence of numbers (integers) -> Sequence of numbers stored as array.
- Samples are controlled by the acquisition (sampling) frequency .
    - e.g., CD quality Fs = 44100 Hz = 44.1 KHz.
    - Phone communications 8000 Hz = 8 KHz.

From Soundwave to Voltage

- **Microphone structure (Condenser):** Sound waves hit Diaphragm Change in capacitance between Diaphragm and Backplate Voltage change.

DIGITAL AUDIO REPRODUCTION

- PC sound card performs the opposite operation.
- Digital signal -> DAC + interpolation -> Continuous time tension -> Speaker (Viceversa of Microphone).

AUTHENTICATION OF DIGITAL AUDIO

- Used in criminal investigations such as bribery, political corruptions, drug deals, and other racketeering activities.
- Evaluated for potential use as digital evidence in court.
- **Goals:**

    Integrity:
    - Detection of discontinuities in the recorded signals as indicators for deletions.
    - Detection of specific attacks that violate the source characteristics (filter, compression, insertion, mixing).

    and

    Authenticity:
    - Authentication of the microphone as the acquisition sensor (device).
    - Authentication of other selected contextual components (recording location, room characteristics, original or playback, background noises).

- **Analyzed features:**
    - Electric network frequency (ENF) of a recording setup.
    - Content-based consistency analyses on local phenomena.
    - Microphone response-based pattern recognition using acquisition setup nonlinearities. You can see the microphone as a filter

ELECTRICAL NETWORK FREQUENCY (ENF) ANALYSIS

Is signal you will find in audio traced and is related to the power supply.

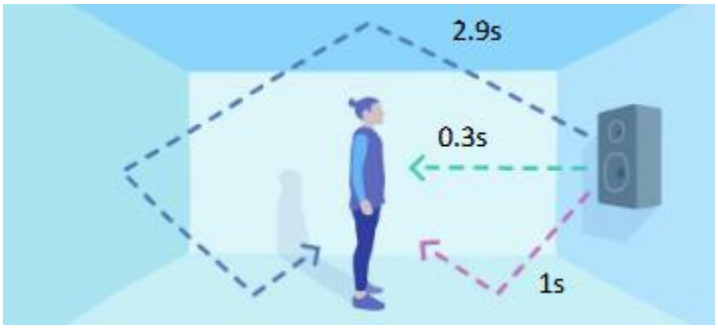
The power signal is not a pure sinusoids, and change depending on the location where you are because the propagation from the power plan is unique.

- In audio recordings, electrical power introduces a background hum; frequency changes in the hum are a fingerprint of a precise location. So I can localize the location where these specific recording was acquire.
- Slight differences are also present for different rooms within a building.
- **Sequence:**
    1. Divide the sequence into overlapping frames.
    2. Compute Fourier Transform (STFT).
    3. Select spectrum for the different harmonics of the nominal frequency f_0 = 50 or 60 Hz.
    4. Create a weighted spectrum: $S(f) = \Sigma_{k=1}^{L} w_k P_{B,k}(kf)$.
    5. Compute features (mean, variance, AR models) and classify using ML.
- A similar signal can be extracted from videos (artificial illumination).
- We can do this also for verify audio deep fake.

SOUND PROPAGATION

- Soundwave propagates through space at a certain speed (it takes some time to hear the sound).
- Speed examples (m/s): Air (20°C): 343; Water: 1482; Steel: 5960.
- Sound can be reflected by objects.
- While propagating, sound loses energy attenuating the amplitude.
- **Attenuation:** Depends on distance and relative humidity/frequency.

REVERBERATION AND EVENT LOCALIZATION



In the example to the subject arrive the original sound and all the replicas (reverberation). The more complex is the room the more complex the reverberation will be. The reverberation leaves data about the place where it was recorded. “This is the impulsive response of the environment”.

- Soundwave propagates towards all the directions.
- Rays (soundwave) are partially reflected and partially absorbed by objects.
- Multiple replicas of the same sound arrive to the listener with different delays: **reverberation**.
- Defines the location of the event in the room.
- Identifying event and reverberation can also be a scene classification strategy.

MICROPHONE’S STRUCTURE

We can have information about the microphone. You can treat your microphone or your audio device as a filter, you can measure the frequency response of the “filter”.

- **Characteristics:** transducer type (dynamic, condenser, electret), sensitivity, frequency response, directionality (polar pattern).
- **Polar patterns:** Omni-directional, Cardioid, Hypercardioid, Supercardioid, Figure-Eight.
- Different microphones (e.g., iPhone versions) have specific frequency response curves.

Feature Extraction

- Used for Microphone Model Generation and Decision.
- **Features:**
  - **LPCC:** linear prediction cepstrum coefficients.
  - **PLP:** perceptually-based linear predictive coefficients.
  - **MFCC:** mel-frequency cepstral coefficients.
- Can distinguish between brands/models (e.g., Sennheiser, Logitech, Sony, etc.).

LECTURE 10

QUALITY ENHANCEMENT

The idea about image and video is related to the fact that often you are ..... permission.

ENHANCING IMAGES/VIDEOS

- Images/videos are usually taken in adversarial conditions (e.g., video surveillance at night).
- Need to process evidences so that quality is enhanced.
- Software or source code can be used.

Softwares:

- **Adobe Photoshop:** Photoshop remains the industry standard for professional photo editing.
- **Adobe Lightroom:** Lightroom is ideal for photographers who manage large volumes of images.
- **Capture One:** Praised for its excellent color management and RAW processing capabilities.
- **Corel PaintShop Pro:** A strong alternative to Photoshop, offering powerful editing tools with a one-time purchase.
- **Affinity Photo:** A cost-effective, single-purchase software with professional-level editing features.

Free and Open-Source:

- **GIMP** (GNU Image Manipulation Program).
- **Darktable and RawTherapee:** These applications are designed for processing RAW files.

DATA ENHANCEMENT – IMAGE/VIDEO

- Tools for image/video enhancement in multimedia forensics:
    - **Sharpening:** Makes edges clearer and more distinct.
    - **Video stabilization:** Reduces the amount of movement in the video.
    - **Distortion correction.** (fundamental if you want to measure something)
    - **Perspective distortion correction.**
    - **Deblurring.**
    - **Noise reduction.**
    - **Masking:** Covers the face or areas of the video that may protect a witness, victim or law enforcement officer.
    - Enhancing contrast, reducing blurring, equalizing color.
    - Magnifying resolution of image.
- **Related to older video technology systems:**
    - **Interlacing:** Interlaced scanning is used in analog systems to record images. A process called de-interlacing may be used to retrieve the information in both fields of video.
    - **Demultiplexing:** Allows for isolation of each camera. In CCTV systems, combines multiple video signals into a single signal or separates a combined signal.

CONTRAST ENHANCEMENT

In this case you change the dynamic range of your pixel. You want to extend this range, so the object is more visible.

- Image pixels are transformed.

$$I^g(x,y) = (I(x,y))^g$$

$$I'(x,y) = (I^g(x,y) - m)/(M - m)$$

- Given parameters g,m,M, we can convert one image into another.

Code: Histogram Equalization & Contrast Enhancement

- **Histogram equalization:**

```
r_image, g_image, b_image = cv2.split(image_src)
r_image_eq = cv2.equalizeHist(r_image)
g_image_eq = cv2.equalizeHist(g_image)
b_image_eq = cv2.equalizeHist(b_image)
image_eq = cv2.merge((r_image_eq, g_image_eq, b_image_eq))
```

- **Contrast enhancement:**

```
image_flattened = image_matrix.flatten()
image_hist = np.zeros(bins)
# histogram creation you can use cv2.calcHist
for pix in image_matrix:
  image_hist[pix] += 1

cum_sum = np.cumsum(image_hist) # cumulative sum
norm = (cum_sum - cum_sum.min()) * 255
n = cum_sum.max() - cum_sum.min()
uniform_norm = norm / n
uniform_norm = uniform_norm.astype('int')
```

- You could use CLAHE (you do not have full invertibility).

SHARPENING

In sharpening you filter the image with highpass filter. Sharpening can be performed by applying a filtering on the image.

```
kernel = np.array([[0, -1, 0], [-1, 5,-1], [0, -1, 0]])
image_sharp = cv2.filter2D(src=image, ddepth=-1, kernel=kernel)
```

MAGNIFICATION

- It is possible to increase the size and the level of clarity of an image.
- When video is involved, it is possible to use multiple frames and do it on motion.

Magnification Code

- Eulerian Video Magnification.
- Or for single image you use magnify.
- Or rescale:

```
pip install PyEVM
from wand.image import Image
Image.magnify()
```

```
# Scaling up
scaleUp = cv2.resize(source, None, fx=scaleX, fy=scaleY, interpolation=cv2.INTER_LINEAR)
```

If you have loss focus on a video you ca compensate blurring. (BLURRING COMPENSATION)

PERSPECTIVE CORRECTION

Very usefull for forensics.

- Code example:

```
matrix = cv2.getPerspectiveTransform(pts1, pts2)
result = cv2.warpPerspective(frame, matrix, (500, 600))
```

NOISE REDUCTION

You have a lot of noise, you can reduce it with classical techniques: adaptive Filtering or Neural network strategies. You preferred to avoid neural network because (ask to gemini) (I think he talks about legal reason).

LENS DISTORTION CORRECTION

- Many cameras use highly-spherical / non-spherical geometry lenses: strong distortion due to curvatures and non-idealities.
- Strong distortion to be corrected.
- You can't do measurement with this.

FRAME INTEGRATION

- Ask to gemini

AUDIO ENHANCEMENT

Microphone acquire information from all the directions, so it can be hidden.

Audio enhancement refers to the process of improving the quality, clarity, and intelligibility of audio signals. Denoising is essential for removing unwanted noise, distortions, or artifacts from audio recordings. Sometimes the forensics looking for noises, because noises can give information about for example the location or the sequence of the event.

- **Noise to be removed/isolated:**
  - **Background:** Persistent, low level due to environment, air cooling, fans, …
  - **Impulse noise:** Clicks, cracks due to electric interference, mic handling, …
  - **Electrical noise:** Hum, buzzes due to electrical devices and recorders, mics, …
  - **Reverberation:** Due to reflections, rooms, …

FORENSIC AUDIO ENHANCEMENT AND ANALYSIS

- **Tools:**
  - *Noise reduction*.
  - *Speaker identification*.
  - *Speech identification. (understand and eventually transcript every word)*
  - *Environment estimation/localization*.
  - *Event reconstruction*.
  - *Authentication*.
- **Examples:** Watergate investigation, 18½-minute gap in an audio recording of President Richard Nixon discussing the Watergate break in with his Chief of Staff. Analysis determined which recorder made the erasure and how many different erasures were made. The level of AC hum recorded to tape even provided details about the acquisition location.

Very frequently, audio is cleaned and amplified and processed by human listening: human brain performs filtering and focusing (be aware of pareidolia)
Automatic detection can be used for monitor and surveillance (count how many vehicles, …)
Tasks (the human do) in evidence processing are:

- Critical listening
- Equalization: select the frequency you want, select where the information is located
- Dynamics processing: expand (highlight) or soften (depress) the components
- Reverb and de-reverb
- Remove or reduce unwanted sounds (air conditioning, electric equipment, …); not continuous sounds are more troublesome (e.g., dog barking) -> *spectrogram analysis*: sounds can be localized and removed
- Electronic measurements and visualization (important especially if you want localize some events)

You can separate voice from music in a song. You can do that because you know the frequency. By creating a mask of filtering that's separate the different part, then when you select the voice and the music "you invert the spectrum".

SPEECH COMPARISON

- Compares a known voice sample to an unknown voice sample. (for identification of people)
- Likelihood of whether the speakers are the same or different.
- Also used to transcript messages.
- Enhancement can be used to highlight whispering and distant voices.
- **Specific Audio Enhancement Examples**
  - Handheld mic in a pocket: AC noise, wall reflections.
  - Wireless mic to squad video recorder: traffic, poor quality.
  - 911 call: Handling noise.
  - Miniature video camera: Distortion reduction.

Example Workflow: Analyzing an Image

Objective

- This relation is intended to improve the quality of an image under analysis.
- In particular, we want to verify if it is possible to:
    - Improve the quality of the image reducing noise and blurring.
    - Increase the detectability of the face of the person depicted in the frame.

Hypothesis

- Taken during night or evening hours (low contrast and range).
- Presence of noise.
- Blurring is affecting the image.
- Resolution is small.

Possible Solution

1. Denoise the image via *fastNlMeansDenoisingColored* from OpenCV2. (Denoising and quality improvements of the picture)
2. Perform some gamma correction.
3. Perform some deblurring with a sharpening filter.
4. Select the detail of face. (tagli dove c'è la faccia, Face extraction)
5. Enlarge the detail by 3 times. (Face magnification)

Analysis Steps (Code)

- **Denoising:**

```
Idenoised = cv2.fastNlMeansDenoisingColored(I, None, 4, 4, 3, 7)

image_eq = (Idenoised - Idenoised.min()) / (Idenoised.max() - Idenoised.min()) * 255

image_eq = image_eq.astype('uint8')
```

- **Gamma Correction:**

```
# gamma via lookup table

invGamma = 1.1

table = np.array([((i / 255.0) ** invGamma) * 255 for i in np.arange(0, 256)]).astype("uint8")

image_eq = cv2.LUT(image_eq, table)
```

- **Sharpening:**

```
kernel = np.array([[0, -1, 0], [-1, 5, -1], [0, -1, 0]])

image_sharp = cv2.filter2D(src=image_eq, ddepth=-1, kernel=kernel)
```

- **Detail Extraction & Resize:**

```
image_detail = image_sharp[350:420, 290:360, :]
width = image_detail.shape[1] * 3
height = image_detail.shape[0] * 3
dim = (width, height)
Ifinal = cv2.resize(image_detail, dim, interpolation=cv2.INTER_LANCZOS4)
```

HOW TO WRITE A REPORT

Structure: The report is usually divided into the following parts:

1. **Synthesis:** define the Objective, Summary of conclusions. 1 page long.
2. **Relation about the evidence analyzed:**
    - What did you get?
    - What did you use (softwares, tools, ...)?
    - What do you hand in? (What is the results?)
    - List of appendixes.
    - ISO/IEC 27037 declaration (reproducibility).
3. **How do you extract the evidence you analyzed**. (gives the details)
4. **Analysis.** (justify everything)
5. **Conclusions.**
6. **Your CV.** (you must prove you are a forensics expert)

SYNTHESIS: OBJECTIVE, ACTIVITIES AND CONCLUSIONS IN SHORT

- **What are the questions that are posed/answered by the relation?**
    - Write clearly what your analysis is intended for.
    - Better to have a style like "hypothesis testing".
    - Define two or three synthetic questions so that the conclusions are clear.
- **What have you done?**
    - Write what activities you carried on.
    - Better a short but clear list.
- **What can you conclude about the evidence?**
    - Answer the questions you posed at the beginning.
    - Be synthetic but clear and detailed.

EXAMPLE

Objective: This relation is intended to improve the quality of an image under analysis. In particular, we want to verify if it is possible to

- Improve the quality of the image reducing noise and blurring
- Increase the detectability of the face of the person depicted in the frame.

Summary of the activities:

As regards the objectives of this consulting, the activity has been carried on as follows:

- Denoising and quality improvements of the picture
- Face extraction
- Face magnification

Conclusions: The operations that we have carried on the frame has allowed to improve the quality of the image. More precisely, we can answer that:

- Image quality can significantly improved
- We can detect a face of a caucasian woman, short haired, whose identity could correspond ...

RELATION Section

- **What did you get? (Received material):**
  - Names, characteristics and details about the original files that you have analyzed/processed (extension, format, frame rate, …).
  - Names of softwares that you have received (e.g. some playing tools in case video sequence has some proprietary format).
  - Better if you report also hash signatures for verification. You must do the hash signatures because is a way to authenticate the evidences, every file must have this.
- **What did you use? (Adopted softwares):**
  - Name all the tools that you have used in the analysis. (for example if you use python specified the version that you use)
  - Better if you use open source solutions.
  - Name also OS and hardware tools (PCs) you have used.
- **What do you hand in? (Produced material):**
  - If you produce a folder with videos, software tools, the report, etc… document all these stuff.
  - Report also a file with a hash signature for all the files.
  - Include also the copy of the relation.

Example of Relation Content

- **Received material:** File name, Format (e.g., PNG), Resolution (e.g., 1024x768), MD5, SHA-1.
- **Adopted softwares:**
  - Laptop specs and OS version.
  - Tools md5 and shasum.
  - Python version and download link.
  - Libraries used (opencv-python, matplotlib, shutil, numpy, mpldatacursor) with versions and links.
  - Script name created for analysis.
- **Handed-in material:** Description of folders (/softwares, /original_image, /created_image), hash files.

Appendixes & Reproducibility

- The relation is made of N pages, it is linked to a digital archive and includes a final appendix.
- **Reproducibility:** The analysis follows the standard for handling digital evidences ISO/IEC 27037 and all the results can be replicated using the included software.

HOW DO YOU EXTRACTED THE EVIDENCE?

Acquisition process

- Report the way you extracted the material (i.e. frames in a video sequence).
- If you select some parts or instants, explain why (e.g. the subject of interest entered the field of view of the camera at …).
- Explain why you used some specific tools and their configuration.

Analysis Section

- Write down what you have done. (all the operation)

- Report all the parameter values and details of your analysis.
- Report intermediate results.
- Be detailed but clear and simple.

Conclusions Section

- State more clearly and finally what you have anticipated.
- Signature and date.

HOW TO AVOID MISTAKES

- **Avoid personal opinions:** Objectivity is key in forensic analysis. Stick to the facts, also in the writing style. Drop phrases like "I believe" or "In my opinion."
- **Motivate all claims:** Describe your methods clearly, explain how evidence supports your conclusions, use data, charts, or images as backup.
- **Report must be consistent:** you must be uniform, otherwise doubts may arise (e.g. no mixed units, different date formats, change of terms).
- **Check data:** Expert review, double check data, track changes and version, be careful to legally relevant details (e.g. names of people or in general personal information).

  (cross-reference data, check calculations, verify sources, review dates and times)

LEGAL ISSUES

- **Validity and authentication:**
  - Clear chain of custody.
  - Come from a qualified expert. (the source has to be qualified)
  - Use scientifically valid methods.
  - Evidence was not tampered. (the hash signature is important)
- **Chain of Custody Table Example:**

What	Who	When	Where
Collection	Officer John Brown	2023-06-15, 10:30 AM	Crime scene
Storage	Evidence clerk	2023-06-15, 2:00 PM	Locker #42
Analysis	Lab tech Sarah Lee	2023-06-16, 9:00 AM	Forensic lab

- **Privacy issues:** Often contain sensitive info. How do you balance openness with privacy?
  - Use initials or case numbers instead of name.
  - Cut out personal info that's not case related.
  - Follow data protection laws for digital evidence.
- **Double checks the information is very important.**
- **Ai is not reliable especially for the quoting of the evidences.** If you report a wrong source (ai generated) you can legally process.

LECTURE 11

Digital Editing of Pictures: Pictures and Information

- We are going to verify if an audio or an image as been edited or not
- Important in fake news detection.
- **Examples of Manipulations:**
  - National Geographic (1982: First example of digitally-altered picture.) - Pyramids moved closer.
  - Time Magazine (1994) - O.J. Simpson's mugshot darkened.
  - Other examples showing people removed or added (Stalin, Mussolini, etc.).
- **Scientists do fakes too** (Example of gel electrophoresis image manipulation). 80% of medical papers include photoshop images. Another common thing in medical research is that instead of upload is own images they copied images from other papers.

IMAGE FORGERY

- **Manipulation:** changing of any image's objects or contents without giving any indication that anything has changed.
- **Image Splicing:** One of the most popular techniques; adding or hiding things by copying portions of one image and pasting them elsewhere in the same image or into a different image.
- **Examples:**
  - 2014: Malaysia Airlines disaster.
  - 2012: Nuclear Tests (missiles cloned).

BASIC PRINCIPLE: HOMOGENEITY

Most of the tools described before permits detecting alterations on images: the fact is that the outputs of these detectors are not the same on all the regions of the image. First you look for uniformity, you can copy a detail from an image and copy on another image, how you can verify that: you compared different features (likes PRNU, CFA, DJPEG) if they different that's mean there was an image splicing. You can do this also in ai images.

CHECK CONSISTENCY/INCONSISTENCY OF FOOTPRINTS

Check consistency of:

- PRNU
- CFA pattern
- Interpolation characteristics
- Local correlation
- Coding parameters
- Double/multiple vs. single compression
- Editing traces
- Physical phenomena (shadows, motion, illumination, ...)
- Geometric inconsistencies
- Similarity/dissimilarity between different parts
- Similarity dissimilarity between different images

Whenever elements expected to be similar are different, raise alarm! We can find this kind of phenomena also in ai images. Often you can verify similarity because you take input from different image to another image.

SIMILARITY ALARM: DUPLICATE ELEMENTS

- Sometimes similarity is a problem: a part of the image is copied and replaced. You replace a part of the image with another part of the same image. So you can search similarity in pixel (it's very strange that in nature a leaf is identical in two different part).
- Useful to hide something
- **How to check?**
    - Scan the image with windows (or SIFT descriptor to find similar points)
    - When points are the same: alarm
    - No way to find which is the original part

Forensically Example

- Tools like "Forensically" allow magnification, histogram equalization, clone detection, error level analysis, noise analysis.

CONSISTENCY FROM PHYSICS AND REAL ENVIRONMENTS

- A different set of strategies checks the physical characteristics of the scene: (we can use also for ai images -> usually are not very physical consistent)
    - Light and illumination
    - Shadows
    - Perspective
    - Reflections
    - Motion (for video)

ENVIRONMENT ESTIMATION

It is possible to estimate the geometry of the environment from audio and video.

- In case it is not consistent, a fake is detected.
- **3D estimation (SfM):** Structure from Motion.
- We estimate the 3D scene and we estimate the reverberation and if the two things are not match that means it's a fake

SHADOWS AND PERSPECTIVE

- Parallel lines obtained by connecting corresponding points in a reflection are projected in lines intersecting in the corresponding vanishing point onto the 2D image. Check for consistency in cast shadows relative to the light source.

INCIDENT LIGHT

- It is possible to mode the direction of the incoming light (under some assumption)
- Including an object in a picture keeping the illumination consistent is quite difficult
- Automatic check algorithms exist.

REFLECTIONS

In authentic images, rays intersect in the vanishing point corresponding to the direction of the projected mirror normal.

The intersection of relaxed constraints: compatible region for the reflection vanishing point. For synthetic image, intersection is not granted.

Video Editing

- Compute the difference between frames and you can find anomaly pixel.
- Detection of removed objects or added elements in video sequences.

BALISTIC MOTION

- After compensating the perspective distortion, it is possible to build a parabolic motion model.
- Tracking the object on video and comparing the trajectories highlight the fake. (e.g., gravity acceleration consistency).

Fake Images from Computer Graphics (CG)

- Comparison between CG and Real images (e.g., faces, landscapes).
- It is becoming increasingly difficult to distinguish CG from real photos.

MULTIPLE-ORIGIN TRACKING

You can trace (online search) images. You can verify that your image is obtain from another images that you find online.

MULTIMEDIA FORGERY AS A CYBERSECURITY ISSUE

- Forged media contents are nowadays very diffused.
- Deep learning has dramatically increased the phenomenon.
- **Legal use:**
    - Video editing and retouching in movie industry.
    - Generation of new products, contents, characters.
    - Artistic and entertainment uses.
    - Assisted design.
- **Illegal use:**
    - Use of deepfakes in propaganda, public opinion manipulation.
    - Creation of revenge porn material, illegal content production.
    - Generation of fake audio files mimicking other people voice for authentication scams (e.g., Fraudsters Used AI to Mimic CEO's Voice).

Deepfake Tools

- **Faceswap:** Using two encoder-decoder pairs.
- **Faceswap-GAN:** Adversarial loss and perceptual loss for auto-encoder.
- **DeepFaceLab:** Expand from the Faceswap method with new models.
- **Avatar Me:** Reconstruct 3D faces from "in-the-wild" images.
- **Neural Voice Puppetry:** Audio-driven facial video synthesis.
- (See full table in slide 23 for more tools).

DEEPFAKE DETECTION

- **GAN Detection:** Look for artifacts.
    - Artifacts on hair/eyes: high frequency details.
    - Face geometry artifacts:
        - Smooth cheeks and forehead.
        - Eyes/eyebrows, shadows.
        - Glasses: is there a glare?
        - Facial hair/mole.
        - Blinking (or lack thereof).
        - Lips' color.
- Note that some of these detections can be overcome with a proper training [e.g., blinking example].
- The Deep fake is continuously evolving.

- **STATISTICAL ANOMALY**
    - GAN images may have different statistics from natural images
    - Benford's Law can capture these traces
    - Given a natural image (JPEG compressed), compute the first digit (FD) of quantized DCT coefficients. It is known that distribution of the FDs follows Benford's law.
- **Summary of Deepfake Detectors:**
    - Eye blinking (LRCN).
    - Intra-frame and temporal inconsistencies (CNN/LSTM).
    - Face warping artifacts (VGG16, ResNet50).
    - Head Poses (SVM).
    - PRNU analysis.
    - Phoneme-viseme mismatches.
    - (See full table in slide 26).

DATA HIDING: WATERMARKING AND STEGANALYSIS

WATERMARK IN HISTORY

- **Examples:**
    - "The Ambassadors" by Hans Holbein (16th century) - Anamorphic skull.
    - Vexierbild (Schoen, 16th century).
    - Cylindrical Anamorphism (18th century).
    - Linguistic Steganography (Boccaccio, Hypnerotomachia Poliphili).
    - Cardan's grid.
    - First system for 'copyright protection': Claude Gellée de Lorraine (ca. 1635).
- **Intuitive Example:** Hiding a message ("ATTACK TOMORROW") using initials of words in a text or filtering. You hide something inside another file.

A TAXONOMY OF WATERMARKING

- **Data Hiding:**
  - **Steganography:** Hiding a message (or some specific content) into an object. Low detectability. Robustness is not important. Only people that know that exactly procedure can access to the information.
  - **Covert Comms.**
  - **Watermarking:** A message adhered to an object, carrying complementary info without altering object semantics. It's like a signature in a file.
    - **Fingerprinting:** (strong watermarking, the signature cannot be erased) use for: Traitor Tracing, Copyright protection.
    - **Reversible wm:** use for: Medical, Military.
    - **(Semi)fragile wm:** (it's designed to be altered if someone does something to the image) use for: Annotation, Authentication
    - **(watermark dna)**

Some Terms

- **Watermark:** It is a message that is adhered to an object, which may carry complementary information and does not alter the object semantics.
- **Visible/Invisible watermarks.** (most of the times I want to use invisible watermarks) (When we talk about invisible watermarks, we do some alteration to the pixel of an image but this alteration is imperceptible, hard to eliminate and the quality is not affective.
- **Imperceptible data hiding:** Harder to eliminate; does not affect the quality of the object to be protected.
- **Steganography:** Hiding a message into an object. Low detectability. Robustness is not important. The different from watermarking is that you **hide** the message.
- **Attack:** Operation that aims to eliminate, distort or extract the hidden data in an unauthorized way. Can be intentional or non-intentional.

WHY NOT CRYPTOGRAPHY?

- Cryptography solves secure distribution but once access is granted, it does not prevent copying.
- Managing access rights: conditional access systems.
- Once access to digital media is granted, cryptography does not prevent from making copies.
- Integrity protection and illegal use identification: watermarking. (you protect the origin of the image)
- Both technologies are complementary, not competing.
- Watermarking provides integrity protection and illegal use identification.
- **Watermarking features:**
  - **Blind** (typically): recovery without the original.
  - **Robust**: Resilient to both intentional and non-intentional attacks.
  - **Imperceptible**: Must take into account human senses.
  - **Security and keys**: Access must be forbidden if there is no access to the keys; key generation, distribution and management.
  - **Conflict resolving**: Must stand multiple watermarking.
  - **Payload**.

WATERMARKING VS. DATA HIDING

Watermarking:

- Embed a key-dependent watermark (maybe unique) and detect its presence
- Binary hypothesis test (yes/no answer).
- Performance: ROC.

Data Hiding:

- Embed multiple bits of information in a key-dependent way and decode them.
- Multiple hypothesis test
- Performance: BER (Bit Error Rate).

TWO FAMILIES OF WATERMARKING

1. **Additive spread-spectrum (Cox, 1996):** Add the watermark to the host signal.
    - Host interference present.
    - Easy to analyze.
2. **Quantization-based (Chen and Wornell, 1999):** Host is quantized with information-dependent quantizer.
    - No host interference.
    - Harder to analyze.

- A combination of both strategies: **STDM** (Spread-Transform Dither Modulation).

A MODEL FOR DATA HIDING

- Information Modulation Perceptual Analysis Embedding Channel (Attacks) Information Extraction. A Perceptual Analysis is the first stage that's allows you to identify those part where you can hide information (specific pixel) without being notice or create a damage. So you have a key (your watermark) and from that you can create additional information. You can generate a single noise and verify if the watermark is presence or not.

HIDING ONE BIT

- Most perceptual experiments concern visibility of noise-like signals and limit its local variance.
- Add/subtract an invisible noise-like watermark to be decoded later.
- so the pixel is *x* and the bit is *b*. You add to x the value b multiply by alpha (you choose alpha)

x: luminance of the pixel in the host image
b: symbol to be embedded $b = \{-1, +1\}$

generate $y = x + b\ \alpha$

where $w = b\ \alpha$ is a watermark. The parameter α defines how much invisible is the watermark.

- How to extract the message? If you know x, you can compute: $\hat{b} = \text{sgn}(y - x)$
- It can be robust to noise.

BLIND CASE

The original image is not kwon, you estimate the average. If the amount of noise is small this is possible otherwise it will not be possible.

If x is not known, it is not possible to recover b

You can think about subtracting the average, i.e.,

$$\hat{b} = \text{sgn}(y - \mu)$$

Unfortunately, SNR is usually small

$$\sigma(y - \mu) >> \alpha^2$$

You need to spread the spectrum!!!

HUMAN CONTRAST SENSITIVITY

The eyes are less sensitive to gradual changes or very quick changes. So high frequency are not relevant for the eyes.

- Contrast sensitivity function (CSF): The human perception system does not respond equally to all spatial frequencies.
- The eye is less sensitive to extremely gradual changes.
- The eye is fairly sensitive to more rapid changes.
- The eye is decreasingly sensitive to yet higher spatial frequencies.
- The perceptual mask tells us how much we can modify each pixel so that each modification is equally visible.
- Let us focus on the ability of edges to conceal information. (e.g. consider a Gaussian gradient filter)

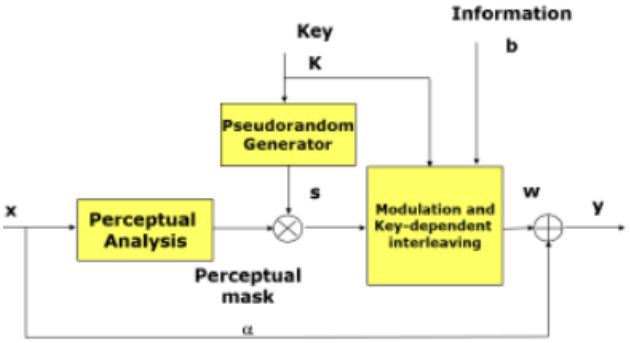
For each pixel, compute horizontal and vertical gradients h[n1; n2] and v[n1; n2].

$$\alpha[n_1,n_2]=\sqrt{h^2[n_1,n_2]+v^2[n_1,n_2]}$$

MODEL FOR SPREAD-SPECTRUM EMBEDDING

You take your image and your divided your image in tiles (called TI, it's basically a small square), each tiles is disjoint. Given this tiles you define an Embedding pass: the combination of signal (pseudorandom generator) multiplies by mask. Multiply this by the bits, so every tiles has one bit. The message is concentrate in a single pixel but is distributed on multiple pixel. This generates the watermark.

- **Embedding:** key-dependent tiling modulated by perceptual mask and pseudo-random sequence.
- **Detection:** Detection is possible without knowing the original image (only key/seed required).



- Assume an image of NxN pixels
- Coordinates $P=\{(n_1,n_2): n_1,n_2 \quad \{1,\ldots,N\}\}$
- Assume a key-dependent tiling

$$T_i: \bigcup_{i=0}^{M-1} T_i = N$$

(each tile T_i is disjoint $|T_i|=\frac{N^2}{M}$

Embedding pulses

$$p_i(n_1,n_2)=\begin{cases}\alpha[n_1,n_2]s[n_1,n_2] & \text{if } (n_1,n_2)\in T_i\\ 0 & \text{otherwise}\end{cases}$$

$\alpha[n_1,n_2]$ Perceptual mask

$s[n_1,n_2]$ Pseud-random sequence with zero-mean and variance 1

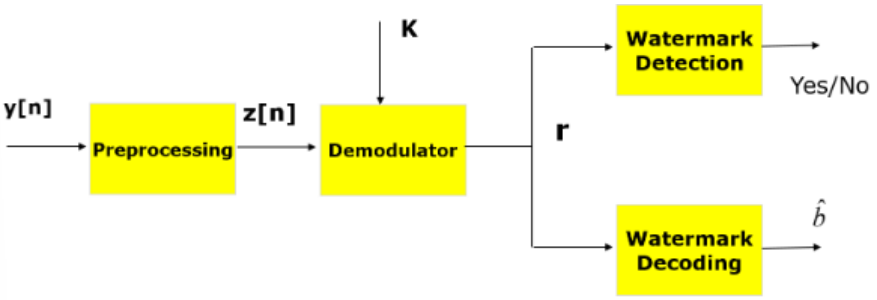
$$w[n_1,n_2]=\sum_{i=0}^{M-1} b_i\, p_i[n_1,n_2]$$ Watermark → $$y[n_1,n_2]=x[n_1,n_2]+w[n_1,n_2]$$

Bits of message

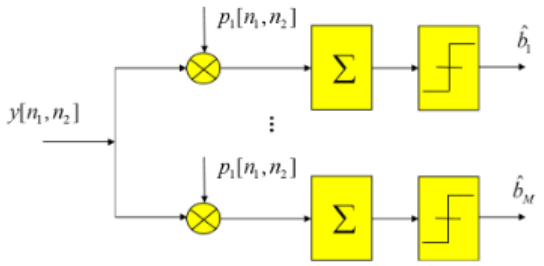
Final image

DETECTION STRATEGY

1.09.00 – 1.11.35 riguardare



Detection is possible without knowing the watermark (of course, the tiling and the pseudorandom sequence must be known).



Let us focus on specific tile (omit i)

$$y = x + \gamma \; b \;\; p$$

If x i.i.d and Gaussian, optimal decoder is easy to find.

$$\hat{b} = \text{sgn} < y, p >$$

Robustness depends on spreading

PERFORMANCES

The shorter the message you want you store the lower will be the error because you use more pixel associate to that specific bit. If you increase the payload the robust goes down as the visibility. If you want to have more robust you reduce the payload and the visibility.

- BER with Gaussian noise attack.
- Trade-off triangle: **Robustness**, **Visibility**, **Payload**.

DCT-DOMAIN WATERMARK EMBEDDING

- Apply the same procedure to DCT coefficients.
- Use Generalized-Gaussian model for different spatial frequencies.

ATTACKS

- Additive noise,
- Filtering,
- Cropping,
- Compression,
- Rotation and scaling,
- Statistical averaging,
- Multiple Watermarking.

APPLICATIONS

- Video/Audio/Software/Text Watermarking.
- Labeling,
- Fingerprinting,
- Authentication,

- Copy and playback control,
- Signaling.

LECTURE 12

Social Media Forensics: INTRODUCTION

Social Media forensics are largely used in investigations activities.

- **Traces:**
  - **The social footprint:** What is the social graph of the user, with whom is he or she connected (friend)?
  - **Communications pattern:** How is the network used for communicating, what method is used, and with whom is the user communicating?
  - **Pictures and videos:** What pictures and videos were uploaded by the user, on which other people's pictures is he or she tagged?
  - **Times of activity:** When is a specific user connected to the social network, when exactly did a specific activity of interest take place?
  - **Apps:** What apps is the user using, what is their purpose, and what information can be inferred in the social context?
- **Issues:**
  - Limited set of data sources in many cases.
  - Acquiring the server's hard drives is not feasible.
  - Leveraging the service operator's data directly requires the service operator's cooperation.

SOCIAL FOOTPRINT

Social graph: it's a graph so it's a mathematical structure, can be oriented or unoriented.

- List of friends can be easily retrieved.
- What type of relations exists within a cluster of friends? This relation connect the various part of the graph. Different data sources can be examined and showed:
  - Social Interaction Graph.
  - Complete Timeline.
  - Localization.
- **Available strategies:**
  - Event tracking (viral scammers, bots, malicious softwares, …).
  - Timeline matching.
  - Differential snapshot. Take a picture of the graph in one moment and see how it's evolved in time. Fundamental because the social graph are non-static.
- **Social Graph Definition:** where are vertices (users) and are relations.
- **Data processed in creating social interaction graphs:**
  - Clustering friends' lists.
  - Graph from picture tags.
  - Direct messages.

DATA ANALYSIS STRATEGIES

- Multi-step processing to identify outliers and group members (e.g., to identify extremists in a network of supporters):
    - **a.** Identify a seed list of known extremists; look for connections and consider the directionality.
    - **b.** Assert cliquishness and in-network focus to improve network membership identification. Fundamental to understand the strategy of people or of the group.
    - **c.** Lexical analysis to identify groups or communities (use of peculiar words or patterns).
    - **d.** Resonance analysis: track propagation in time.
        - Build a linguistic model.
        - Regional matches.
        - Score users with linguistic model; check whether it happened by chance.
        - Quantify matching (likelihood) in time.
    - **e.** Stance analysis: detecting messaging strategies.
    - **f.** Crowdsourcing to understand the information environment (Geotagging, Image classification, Mapping images). Gives you information where you are.

RESULTS FROM SN (Social Network) DATA PROCESSING

- **Understanding Networks:**
    - Detect nuances in the dynamics of interpersonal networks by analyzing the information posted by users on these platforms. (What do people do typically? They analyze the patterns of the news to create very good fake news, try to push your emotional reaction)
    - Regional trends.
    - Identify key roles.
- **Verify credibility of SN data:**
    - Crowdsource verification of developments and events, improving situational awareness for security forces.
    - Sentiment analysis
    - Intelligence
- Detect and counter propaganda on SN.
- Deduce or extract sensitive data.
- Use data for civil and military information. (used social media datas for emergency events)

ISSUES with Social Media Forensics

- Data are hardly reproducible: timeline and graphs can look quite different in time.
    - Data volume is quite big.
    - Social interactions are dynamic (data can be concealed). (for this you do snapshot)
- Some data are available only to ISP or SN managers.
- How deep must analysis go in the interaction graph?
- **Sampling bias:** users share what they want to share; age; social classes.
- Social media penetration is variable (and depends on SN).
- Legal issues and Ethics. (if your friend is a terrorist suspect they investigate also you)

SPREADING OF DATA THROUGH NETWORKS

- Social media and data sharing platforms allow pervasive diffusion of digital contents (Posts, datasets, images, videos, software). A content is propagating throw the SN, when is viral the problem you don't know which is the original one (you find multiple copy, but you don't kwon the history).
- After being available online, contents start spreading (Epidemic diffusion models). (this model where use also to malware detection, the strategy is similar)
- Often contents evolve and change.
- At the end, several copies (or near duplicates) of the same content are available.

Phylogenetic Analysis

- The term was mutuated from biology given the analogy with the analysis of the mutation process that occur to living organisms.
- This evolution can be represented by a tree structure named **phylogenetic tree** (it's a topological tree from a mathematical view). (you can identify some common ancestor)
- It can be done for different media and digital contents.
- **Applications:**
    - Traceback of illegal contents.
    - Copyright protection.
    - Malware profiling.
    - Fake news detection and debunking.

IMAGE PHYLOGENY TREE RECONSTRUCTION

How to reconstruct a phylogeny tree? Graph-based spanning tree algorithms. You have to complete the graph that requires that you have similarity networks.

- **Steps:**
    1. Compute a dissimilarity matrix between all the possible couples of images in the set.
    2. Build a relational graph where edge weights correspond to dissimilarity measurements.
    3. Run **minimum spanning tree** algorithms (if you are using dissimilarity method) on the graph (e.g., Oriented Kruskal or Optimum Branching). (if you are using similarity method you are going to use **maximize spanning tree**)
    4. Obtain a tree that signals which images are parents and which are children.
- **Problem:** Dissimilarity is an extremely noisy metric!

STATE-OF-THE-ART STRATEGIES for Phylogeny

- Improve dissimilarity estimation.
- Combine multiple dissimilarity metrics.
- Probabilistic modelling.
- Dependency checks.
- Use side information.
- Employ new DL-based robust metrics.
- Deep learning can be used to denoise dissimilarity and reconstruct the tree at the same time.

In our approach, deep learning can be used to denoise dissimilarity and reconstruct the tree at the same time.

IoT Forensics

INTERNET-OF-THINGS (IoT) / INTERNET-OF-EVERYTHING

- IoT is the interconnection via Internet or other communication networks of heterogeneous devices enabled with sensors, processing ability, softwares.
- IoT devices are more and more diffuse. In the industrial sector and in the private sector.
- **Examples:**
  - Smartphones and related appliances.
  - Smart homes (domestic automation).
  - Smart cities (traffic, infrastructure).
  - Power grids and energy management.
  - Manufacturing (Industry 4.0).
  - Healthcare systems.
  - Environmental monitoring and Agriculture.
- Communication types: **Thing-to-Thing**, **Person-to-Thing**, **Person-to-Person**.

DEVICES and Market

- Wearable devices (smart watches, glasses, health sensors).
- Smart home appliances (locks, light sensors).
- Control and safety systems for industrial application.
- Vehicles (Cars, UAVs, etc.).
- Market size: ~1.7 trillion $ (in 2020).

NEED FOR IOT FORENSICS

- **Many attack ports/points:**
  - Like having a house with too many windows/doors (extensive attack surface). (more points where malicious attackers can attack)
  - Simplistic devices: Security level is often not adequate (low power/cost). (very low level of protection, weak points of your network)
  - Gateway for malware: IoT devices act as doors to private networks.
  - Heterogeneity: Different devices require different security measures. (every device has its own software so you have to update the security one by one)
- **New cybersecurity threats:**
  - Related to health, privacy, personal goods;
  - crossing boundary into human life threats.
  - **Main issues:** DoS (Denial of Service) and Information leakage.
- **Forensic value:** Can help investigation and prove/disprove hypothesis.
- **Additional source of information:**
  - can help investigation
  - can prove or disprove a hypothesis for digital investigators

CHALLENGES

- **General problems:**
  - Evidence identification, collection and preservation.
  - Analysis and correlation.
  - Presentation.
- **Specific problems:**
  - Lack of well-defined and tested methodologies.
  - Lack of tools. (lack specific tools that's deal which such complexity)
  - This could result in:
    - Risk of contaminating or destroying evidences.
    - Jeopardize trust in cross-jurisdictional activities.

EVIDENCE IDENTIFICATION, COLLECTION AND PRESERVATION

- Detecting presence of IoT system that could be useful in the investigation.
- Lack of training for DEFR (Digital Evidence First Responder).
- Wide range of software and hardware specifications.
- **Lifespan limitations:** Data can be easily overwritten or lost. (this because this devices are very simple, the storage is loss. Most of the data is store on the cloud – it's become from an hardware forensics problem to a cloud forensics problem)
- Officers could shut down or disconnect the device prior to saving necessary information.

EVIDENCE ANALYSIS AND CORRELATION

- Overwhelming amount of data.
- Time lining and limited correlation evidence.
- No metadata. (because it's too simple)
- **Result:** Creating a timeline and handling huge amounts of data is difficult.

EVIDENCE PRESENTATION

- Jury has a basic understanding of digital forensics and cloud computing.
- Explaining technicalities could be very long and challenging.
- No standardization, so methodology might not be accepted.
- **Risk:** Unusual methodology can be challenged in court. (risk of removing or altering part of the data must be refuted)

THREE LEVEL OF ANALYSIS

Tree level of analysis, depending on specific case you could be in one of them or more them:

1. Cloud Forensics.
2. Network Forensics.
3. Device Forensics.

Comparison: TRADITIONAL/CLOUD/IOT FORENSICS

More devices in IoT, also networks are typically RFID, in traditional forensics there are a lot of networks. Different types of boundaries, in traditional forensics the boundaries are more strong, in IoT are almost boundaryless. Everything scale up when we talk about IoT forensics.

Item	Traditional Forensics	Cloud Forensics	IoT Forensics
Number of devices	Billions of devices	Billions of virtual machines and servers and other cloud datacentre components.	It is expected to have 50 billion by 2020 and it is increasing.
Type of Networks	Mobile communication networks, Internet, Bluetooth, Wi-Fi, wired, and wireless networks	Private, public, community and hybrid	RFID and wireless sensor networks (WSNs) such as reader sensor
Protocols	Different types of wireless networks, Ethernet, Bluetooth, IPv4, and IPv6	MTP, EIGRP, AMQP, XMPP, CEE, SSHP, IGMP, SRP, CLNP, and Gossip	RFID and Rime
Network boundaries	Fairly clear boundaries and lines of ownership	Chains of datacentres	Increasingly and there are no defined boundaries
Ownership	Individuals, groups, companies, governments, etc.	Individuals, Cloud Service Providers (CSPs), organizations, governments, etc.	Individuals, groups, companies, governments, etc.
Source of evidence	PC, social networks, authentication, authorization, and accounting servers, mobile devices, and web clients	Virtual machines, virtual servers, network logs, and cloud storage	Home appliances, connected cars, tags, readers, sensors nodes, WSNs, and medical devices.
Type of Evidence	Electronic documents and standard file formats such as JPEG, MP3, etc.	All formats	All formats
Quantity of evidence data	Up to terabytes of data	Petabytes of data	Up to exabytes of data
What to seize	Seize devices as required	Remote access to evidence such as Cloud servers, virtual machines, virtual server (no physical accessibility)	Identify the next best things for the source of evidence

Case Studies (Incidents)

- **Incident 1: TRENDnet (2013):** IP cameras with poor security allowed hackers to access 700 feeds (surveillance of private homes, infants). FTC complaint filed.
- **Incident 2: Volkswagen Scandal (2015):** Emission cheating software detected by WVU researchers during road tests. In-house testing passed, but road test failed.
- **Incident 3: Strava Case:** Fitness app heatmaps revealed US military bases in Syria and Afghanistan due to soldiers tracking their runs.
- **Incident 4: Arkansas Case (2015):** Amazon Echo data requested in a murder investigation. Issue of privacy (4th Amendment) vs probable cause.

IoT Forensics Challenges Summary

1. Diversity of IoT environment.
2. Complexity of IoT architecture.
3. Big IoT data.
4. Multi-jurisdiction.
5. Privacy concerns.
6. Chain of custody and data integrity.
7. Security.
8. Evidence found through IoT forensics.

EXISTING SOLUTIONS

- **Cyber-Trust:** use Intrusion Detection Systems (IDS) with distributed ledger technology (DLT), for evidence collection and preservation within a smart home environment.
- **FEMS (Forensics Edge Management System):** solution for smart home, combines several security functions; difficult to implement and test (for the eternity of the networks).
- **PROFIT Model:** integrates the (ISO/IEC 29100:2011) standard to ensure and maintain the privacy of witnesses' personal data to encourage voluntary participation.
- **Probe-IoT Framework:** distributed digital ledger to maintain track of all transactions taking place between IoT devices, users, and cloud services.
- We need to prove that this type of solutions are valid in order to perform digital forensic in a right way.

ANOMALY DETECTION

Very useful when you want to identify something strange (attack, malfunction…)

WHAT ARE ANOMALIES?

- **Definition:** An anomaly is an object that is notably different from the majority of the remaining objects.
- **Synonyms:** Unusual, irregular, atypical, inconsistent, unexpected, rare, erroneous, faulty, fraudulent, malicious, strange.
- Also called outliers or abnormalities vs. normal or nominal data
- Different in terms of pattern compared to nominal data. (nominal data is used by mathematics)
- **Uses:** Fault detection, Cybersecurity, Medical analysis, Machine vision, Statistics, Law enforcement, financial fraud.

NOISE ≠ ANOMALY

- **Noise** is erroneous, random values or contaminating objects (e.g., typo in weight, contamination).
- **Noise** interesting. (because not reveal something new)
- **Noise** is not an anomaly is an error.
- **Noise** doesn't necessarily produce unusual values or objects.
- **Anomaly:** May be interesting if not a result of noise. (anomaly gives something new)
- Noise and anomalies are distinct concepts.

HYSTORY FACTS: OZONE DEPLETION

- In 1985 ground sensors detected low ozone levels.
- Satellite instruments (Nimbus 7) didn't record them because the low values were treated as outliers/noise and automatically discarded by the software.
- They confused as an error but it was an anomaly.

URGING ISSUES FOR ANOMALY DETECTION

Key technological factors:

- Increased data volume (huge amount data to process):
    - Complex networks (sensors, IoT, heterogeneous terminals)
    - Lots of user generated data

    - Mobile devices with increased data generation capacity
- Data are exchanged and mutate at high speed (you need to update data in time)
- **Consequences:** Severe consequences for failure.
- High failure consequences -> low failure rate (you prefer to have a false positive (false alarm) than a false negative)

Detection side:

- Very accurate detectors with machine learning and deep learning strategies
- Data processing capabilities are improving everyday (we can also work with small device)
- Ability to deal with high-dimensional spaces and large data sets
- Low failure consequences -> Not-so-low failure rate

EXAMPLE OF ANOMALIES

- Network attack (anomalous traffic)
- Image tampering or deepfake detection (local pixel characteristics different from the other ones)
- Fraud detection
- OS attacks and access control analysis
- Malware detection
- not only for forensics:
    - Production chain optimization
    - Failure monitoring and prediction
    - Predictive maintenance

WHY ANOMALY DETECTION?

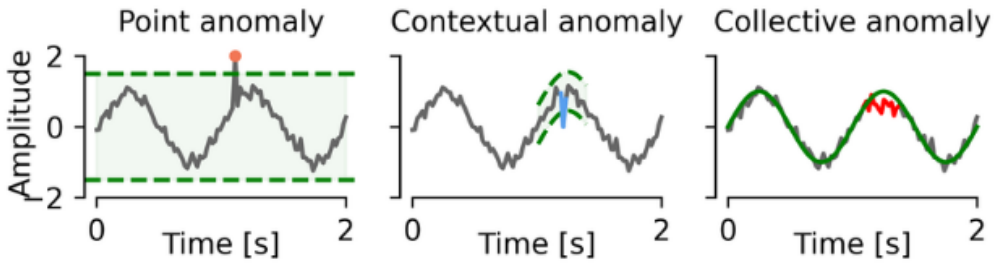
- **Unknownness:** features and characteristics remain unknown until anomalies actually occur.
- **Heterogeneous (and evolving) anomaly classes:** Anomalies are irregular and changing (adversarial attackes) and thus, one class of anomalies may demonstrate completely different abnormal characteristics. (if the attaker change the strategy you have to also update)
- **Rarity and class imbalance:** Rare data events; a large amount of labeled abnormal instances are difficult (if not impossible). (Lots of nominal data where everything is working and you have only few anomalous data, in some training it's possible you don't have anomalous data)
- **Diverse types of anomalies:** Three different types have been explored
- **Need to automatize the process**

THEORY: WHAT IS AN ANOMALY?

- Define a set of data. $x_1, x_2, \ldots, x_n \in \mathbb{R}^d$
- This data is a mixture of "nominal" and "anomalous" samples.
- Anomalies are created by a different generation process with respect with the nominal samples.
- **WDAD (Well-Defined Anomaly Distribution) assumption:** Anomalies drawn from a known distribution. This assumption says that there are nominal data that they are generated by some specific distribution (statistics) and there are anomalous data that are generated by different distributions (statistics).
- **WDAD is not valid if:**
    - Adversarial conditions (fraud, insider threats, cyber security)

- diverse set of modes (new failures, not known)
- user's notion of anomaly changes over time (anomaly = interesting point or new data type) An anomalous event it's incapsulate into the dataset and then you search for a different type of anomaly after this. If the attaker change the strategy you have to change the distribution because the distribution is not more well defined.

TAXONOMY OF ANOMALY DETECTION (three different types of anomaly)



1. **Point Anomalies:** An observation that deviates from trend (e.g., packet loss, single credit card transaction).
2. **Contextual Anomalies:** Violation of a seasonal or contextual trend (e.g., heavy rainfall in London is normal, in Sahara is anomalous; high expenses during Christmas are normal).
3. **Collective Anomalies:** Deviation from a pattern (sequence) relative to the full dataset (e.g., heart skipping a beat, traffic volume spikes between routers). You have an anomaly, if you consider the all dataset it is strange but if you consider the specific value it is not strange.