Investigar y probar el delito en la sociedad de la información

Basic Tools in Forensic Informatics

Cesare Maioli

CIRSFID and Faculty of Law
University of Bologna

Burgos, *October 7, 2011*

Introduction

- Digital information is the form in which the information system carries out its activities, through software applications, emails, data feeds or the Web
- Global economy is increasingly dependent on information processing or transmission of data across networks
- Law enforcement agencies investigate and prosecute the perpetrators of cybercrimes driven by the availability and accessibility of data and information to the investigator, whether in the context of intelligence gathering, evidential retrieval and subsequent analysis and presentation to courts

Contents

- Cybercrime and forensic informatics
- A frame of problems in forensic informatics
- Digital evidence in forensic informatics

Computer-related crime

- Traditional type of criminal offences that may be committed using computers as the instruments of the crime, such as fraud or forgery and more generally inappropriate and misuse of information
- Content-related offences concerned with the use of ICT to facilitate the distribution of unlawful contents or illegal data; examples are criminal copyright infringements and child pornography
- Computer integrity offences which address activities that attack the integrity of computer and communication systems, such as distributing computer virus

Investigative techniques

Law enforcement investigation techniques are subdivided into

- covert techniques, such as interception and surveillance, generally used at an earlier stage of the investigation for the gathering of intelligence as much as evidence
- coercive techniques, such as search and seizure, used primarily to gather evidence once the relevant ICT resources have been identified

Both categories are involved in the investigations on cybercrime

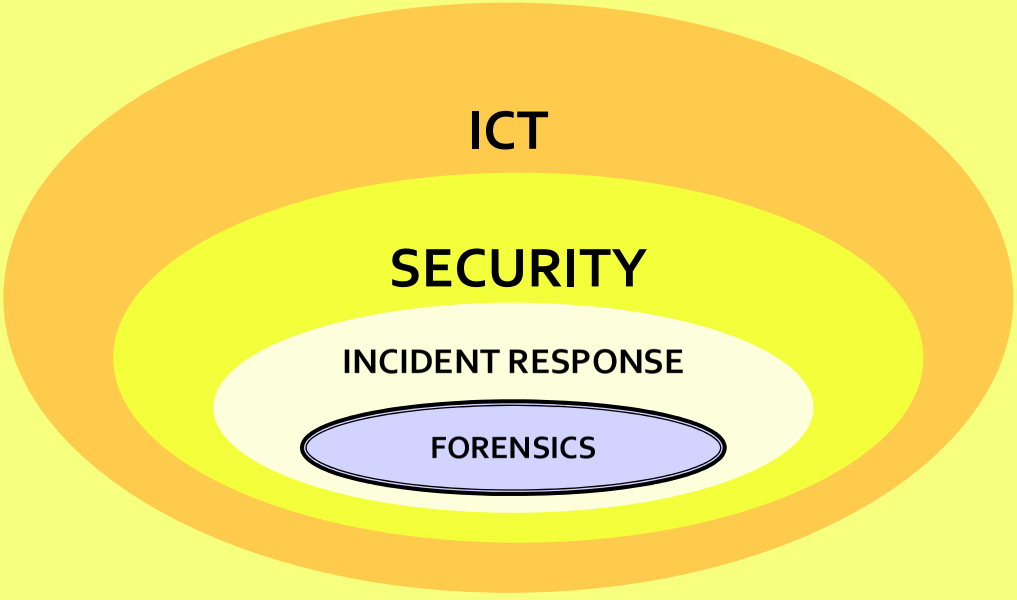
Digital forensics - I

- It involves the preservation, identification, extraction, documentation, and interpretation of computer media for evidentiary and cause analysis
- Digital forensics, forensic computing, cyber-forensics are based on the intangible and often transient nature of data, especially in a network environment
- A process of applying scientific and analytical techniques to computers, networks, digital devices, and files to discover or recover admissible evidence
- Types
  - Disk Forensics
  - Network Forensics
  - Email Forensics
  - Internet Forensics
  - Portable Device Forensics (e.g. flash cards, cell phones, IM devices)

Digital forensics - II

- The technology renders the process of investigation and recording of data for evidence extremely vulnerable to defence claims of errors, technical malfunction, prejudicial interference or fabrication
- The lack of adequate training exacerbate these difficulties
- The practice has developed in a casual manner, based on expediency and circumstances

Context



Issues of network forensics

- The growth of network-based crime has raised some unique and difficult issues in respect of the balance between :
  - the needs of investigators and prosecutors
  - the right of the network users to privacy
- The interests of Communication Service Providers where law enforcements agencies are looking for assistance in terms of
  - gathering data transmitted by the suspects
  - providing data generated by the Communication Service Providers about the suspects' activities
- The main ways to obtaining forensic data are:
  - data from the suspect, obtained covertly, through various forms of surveillance
  - data obtained from a Communication Service Provider
  - data from the suspect, obtained coercively, through the exercise of search and seizure powers

Example: steps in the management of forensic data

- Acquisition
  - Physically or remotely obtaining possession of the computer, all network mappings from the system, and external physical storage devices
- Identification
  - This step involves identifying what data could be recovered and electronically retrieving it by running various computer forensic tools and software suites
- Evaluation
  - Evaluating the information recovered to determine if and how it could be used again the suspect for acquittal or prosecution in court
- Presentation
  - This step involves the presentation of evidence discovered in a manner which is understood by lawyers, non-technically staff, and suitable as evidence as determined by national laws

Contents

- Cybercrime and forensic informatics
- A frame of problems in forensic informatics
- Digital evidence in forensic informatics

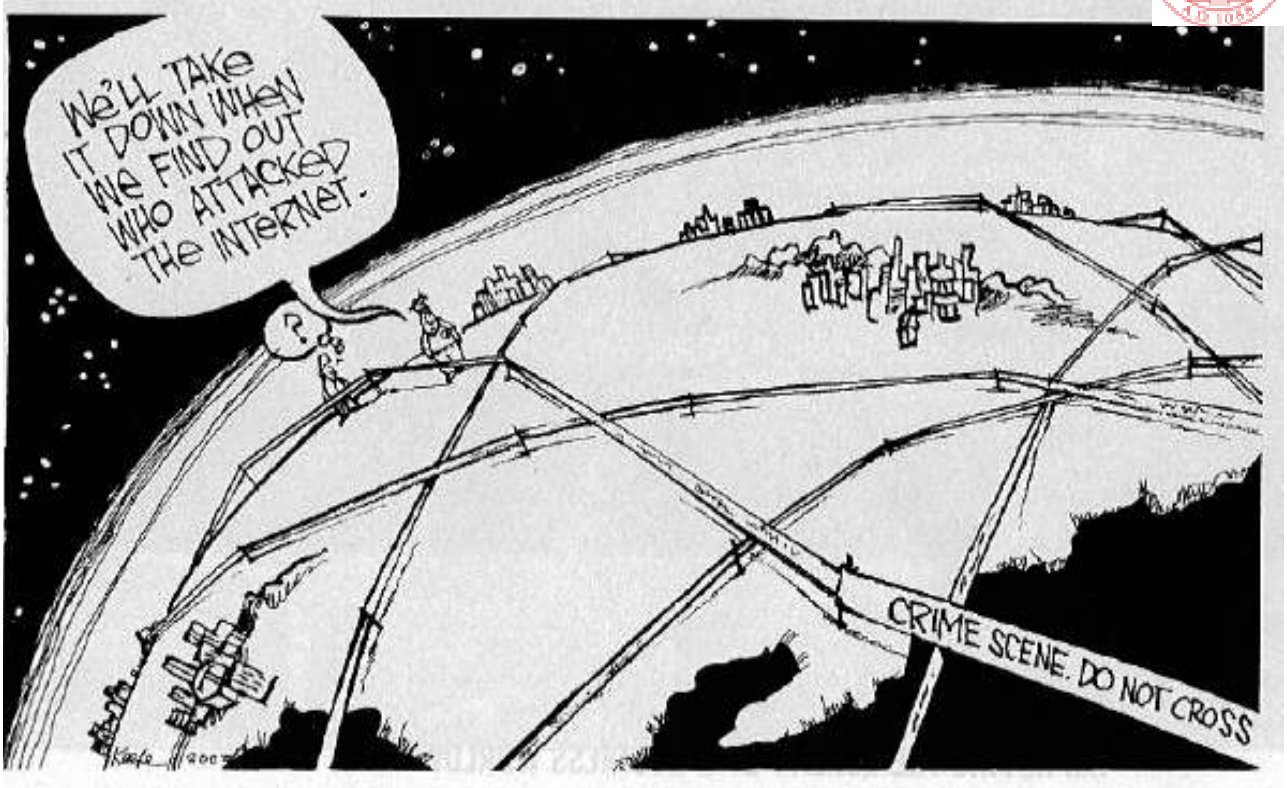
Identity problem

- To establish an adequate forensic link between an item of data and the virtual identity of the person
- To establish an adequate forensic link between the virtual identity and a real person

Location problem

- To identify the physical location of a suspect
- To consider the jurisdictional implications since cyberspace activities cross state borders
- To distinguish between data at rest and data in transmission: the legal distinction between searching computer systems and seizing data stored therein and intercepting data in the course of transmission should be clearly delineated and applied

Frontiers in cyberspace...



http://digitalforensics.champlain.edu/about_cdf.html

Integrity problem

- The process of obtaining forensic data is a significant technical challenge for investigators, since it may modify the source data or its related meta-data (e.g. date and time), thus undermining the evidential value of the forensic material
- The manner in which an investigator interacts with data creates further problems, due to the confused treatment of intangible information under many national law

Stickiness problem

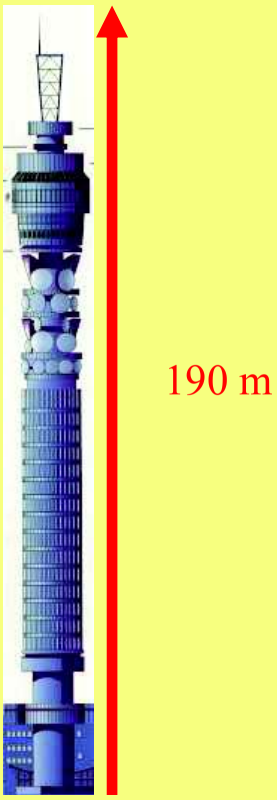
- Multiple copies generated by the communication process
- The manner in which data is held and removed on electronic storage media
- In general stickiness of data work in favour of the investigator
- Conversely, the perception that data held on ICT sources is transient may work in the advantage of the defendant, where he can raise doubt as to the existence of relevant forensic data

Analysis problem

- The volume and nature of the data an investigator may handle in the course of an investigation
- Data media are capable of storing vast amount of data and networks are capable of transmitting huge bit streams of data
- Obtaining and preserving those data are straightforward and rapid
- The ability to access, manage and analyse it for subsequent presentation in the court present problems of overcoming protection mechanisms, respecting any time and budget limit imposed by law

Example: Volume of data

- The BT Tower is 190 m tall
- Printing out a 6 Giga Bytes hard disk drive produces a stack of paper taller than the tower
- A 300 pages book digitally takes about 650 Kilo Byte
- An extension of 10 Giga byte hosts about 15.250 books



Data type problem

- Digital evidence may comprise：
  - the content of a transmission
  - the attributes or meta-data of a communication activity
  - the right of the network users to privacy
  - their management by an ICT resource
- The digitization of information means that traditionally distinct forms of data are represented in a common format: binary machine code
- Existing legal rules treat the obtaining of different types of forensic data differently; e.g. the law governing interception being only applicable, or not, to obtaining communication contents
- Distinguish different types of data and determining the applicable procedural regime governing the obtaining of such data is a challenge

Traceability issues

- Sources
  - various data elements that a suspect has disclosed in the course of his activities, which can be traced back to the suspect
  - data created by a suspect's use of a communication service
  - the contents of a person's communication activities
- Identify the source and the recipient based on some form of unique identities
- Where the terminal equipment resides in an environment where multiple persons have access, proving that a specific person was using the equipment or account at the relevant time, is a problem and potential defence

Example: Resolution of the IP address

- Investigator resolve the IP address of an individual
  1. Identify the IP address that needs to be resolved to a machine or person
  2. Establish an entity, such as Communication Service Providers, to whom the IP address has been assigned
  3. Approach the IP holder to match the target IP to a special user
  4. Obtain the subscribers' account details from the Communication Service Providers' administration records

- The ability to trace a person from his IP address is dependent on the input of different entities and the existence of various logs and records

- Importance of data retention requirements on Communication Service Providers

Contents

- Cybercrime and forensic informatics

- A frame of problems in forensic informatics

- Digital evidence in forensic informatics

Digital evidences

- The distinctive aspect of computer related evidence is that it is volatile, require interpretation before it is intelligible, and may be copied rapidly with absolute accuracy
- The alteration can come about through use by an operator, or by normal processing activities of the computer; it is difficult to determine what changes may have been made, and at what time those changes were made
- Examining digital evidence can be lengthy; thus investigators must be both thorough and cautious when gathering evidence. Usually a pristine copy is quickly produced for examination and the storage drive returned to the application access

Basic principles in digital evidence handling

- The storage device has to be "frozen" in time: the evidence has to be collected as early as possible and without altering the contents of the storage device
- There must be continuity of evidence: there must be an unbroken chain of custody for the evidence from when it was originally collected to where it appear in court
- All procedures used in examination should be auditable: a qualified expert appointed by the other side should be able to track all investigations and repeat the processes carried out by the prosecutor's experts

Example: Chain of custody

- Establishes each person who has had custody of the evidence
- Establishes continuity of possession
- Proof of integrity of the handling of the evidence collected

1. Date and time item was seized
2. Location and who it was obtained from
3. Make, model, and serial number
4. Name of individuals who collected evidence
5. Description of evidence
6. Full name and signature of person receiving evidence
7. Case number and item (tag) number of evidence
8. Hash values (if available) of evidence if able to obtain
9. Pertinent technical data (e.g. drive geometry)

Media functionalities

- Most media have a basic four-way functionality, in terms of enabling data to be written to, read, modified and deleted
- Some media have been developed to only enable data to be written or read; they have been designed to address the integrity problem for the users and they are beneficial from an investigative perspective

Storage media

- Digital media store data in different ways at:
  - the physical level, such has magnetic particles and laser-created pits
  - the logical level, in terms of partitions, drives and sectors
- The way in which a device logically handles the data has direct implications for any subsequent forensic analysis
- The various file systems utilize space on storage media in very different ways, which require the use of different analytical techniques to examine data stored by them
- Under all file systems data is not necessarily stored in continuous manner but rather will often be fragmented across the media, in blocks which are only logically associated through addressing information

Data deletion

- The deletion of data from digital media may take different forms:
  - file deletion in a standard desktop applications only results in the removal of the addressing information associated with each block of data, which logically links the various blocks comprising the content of the file
  - the files that are deleted are renamed in another directory (e.g. Trash folder)
- The data remains on the media, and is partially recoverable, until it has been completely overwritten by new data, or been deleted by other mean (e.g. wiping software)
- The residual physical representation of erased data is referred to as data permanence, and is one of the causes of the stickiness data problem

Identification and organization of operational data

- System software and application software use system for identifying and organizing the data they operate upon, in terms of file names, extensions, folders and directories
- Such systems often maintain a wide range of valuable forensic details about he attributes of a data file, e.g. size and usage
- Usage data, such as the time and date at which an action was carried out on a data file, is an extremely valuable forensic source, but it is also highly vulnerable to claim of inaccuracy, modification and interference
- Thus there is the need of some form of corroboration using other data sources (e.g. data placed in a digital photography)

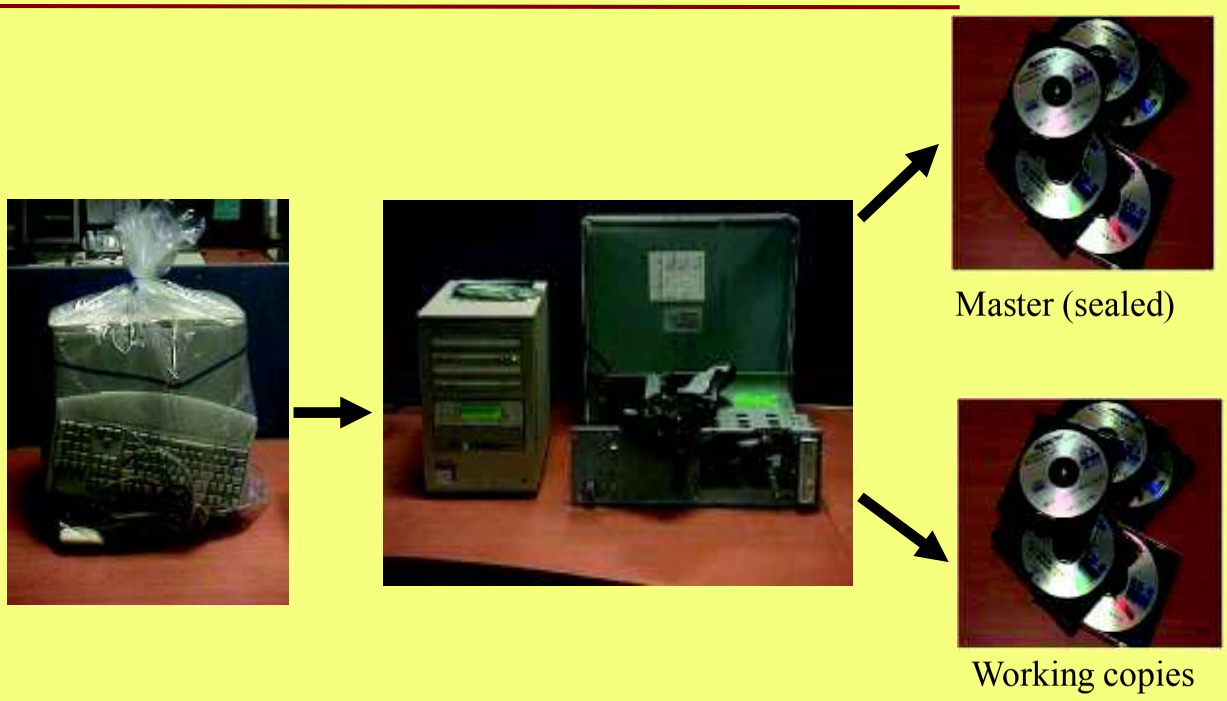
Example: Collecting evidences

- Make exact copies of all hard drives and disks using computer software
  - ⇨ Date and Time stamped on each file; used for timeline
- Protect the computer system
  - ⇨ Avoid deletion, damage, viruses and corruption
- Discover files
  - ⇨ Normal Files
  - ⇨ Deleted Files
  - ⇨ Password Protected Files
  - ⇨ Hidden Files
  - ⇨ Encrypted Files
- Reveal all contents of hidden files used by application and operating system
- Access contents of password protected files if legally able to do so
- Analyze data
- Print out analysis
  - ⇨ Computer System
  - ⇨ All Files and data
  - ⇨ Overall opinion
- Provide expert consultation/testimony

Imaging - I

- Investigators addressing the integrity problem need to be able to obtain the data in a manner which is complete, yet with minimal interference with the target data
- Such data may be printed and copied, although this may result in alterations to the meta-data associated with the target data, which may create evidential vulnerabilities
- Thus the most used technique to obtain forensic data is that of imaging
- A bit stream image of the digital media, e.g. hard disk and smart card, is acquired and created in an non-invasive manner including those areas which are not occupied by data elements

Example: Imaging process



Imaging - II

- This process generates data, such as the cryptographic hash function, that may be required at a later stage to verify the authenticity and integrity of the acquisition process and any subsequent copies
- A number of copies are generated: a master copy, a working copy and copies for the parties in the trial
- Imaging can allow the original digital media to be returned to the owner who can then continue to use the resource
- The images are widely accepted as an accurate representation of the original digital media in courts

Example: Digest and hash function - I

- The digest of a file (sequence of bits) is a predefined fixed length string of symbols generated by the execution of hashing function over the file itself
- DPCM 8 febbraio 1999: *"l'impronta di una sequenza di simboli binari è una sequenza di simboli binari di lunghezza predefinita generata mediante l'applicazione alla prima di un'opportuna funzione di hash"*
- It is impossible from the digest to reconstruct the original text
- Collisions of the same digest value from two different sources are impossible

Example: Digest and hash function - II

I agree
efcc61c1c03db8d8ea8569545c073c814a0ed755

My place of birth is Ravenna
fe1188eecd44ee23e13c4b6655edc8cd5cdb6f25

<*this presentation*>
0e6d7d56c4520756f59235b6ae981cdb5f9820a0

Nel mezzo del cammin di nostra vita... l'altre stelle
a45ba8904e6c531aca987de002e2ef130cc42134

I am an Engineer
ea0ae29b3b2c20fc018aaca45c3746a057b893e7

I am an Engineer.
01f1d8abd9c2e6130870842055d97d315dff1ea3

- Same Length: 40 digits (160 bits -8 bytes)
- They are document content dependent and hash function dependent

TimeLine

- The capability to map the events recorded on the various devices from which data is derived to real time is a critical element in the forensic analysis
- The establishment of an accurate chronology of the events related to the criminal conduct
- Most of the ICTs that operate within a network and at its edges time record; yet this multiplicity coupled with possible inaccuracy makes mapping events in cyberspace a considerable forensic challenge
- The problem is exacerbated by the transnational context and different time-zones in which the networks operate

Convention on Cybercrime

ICT and cyber crime have an obvious international dimension and governments have recognized the need to ensure that legal protection is harmonized among nations

Aims

- Harmonising the domestic criminal substantive law elements of offences and connected provisions in the area of cyber-crime
- Providing for domestic criminal procedural law powers necessary for the investigation and prosecution of such offences as well as other offences committed by means of a computer system or evidence in relation to which is in electronic form
- Setting up a fast and effective regime of international co-operation

Conclusions

- Digitalization makes difficult to ensure that the different types of information obtained from different locations continue to be subject to different legal treatment
- The deployment of evidence gathering within an ICT environment can challenge traditional procedural concepts
- The needs of law enforcement have to be balanced against the rights of the perpetrator, the victim or others caught up in the investigation process
- Obtaining access to data, acquired by the process of seizing systems and data, and analyzing them within the time and resource constraints imposed on the public policies present a significant challenge

References

- Walden I., *Computer crimes and digital investigations*, Oxford University Press, 2007

- Brenner S., *Cybercrime: criminal threats from cyberspace,* Praeger, 2010
- Casey E. (ed.), *Handbook of computer crime investigation: forensic tools and technology*, Academic Press, 2007
- Clough J., *Principles of cybercrime,* Cambridge University Press, 2010
- Lloyd I., *Information technology law*, Oxford University Press, 2011
- Wall D. S. (ed.), *Crime and deviance in cyberspace*, Aldershot, 2009

- http://www.cirsfid.unibo.it/CIRSFID/Centro/AreeDisciplinari/Informatica_Forense.htm