Network Security

MERSTel/MEI – 1st year / 1st Semester

Access Control and Authentication / Controlo de Acesso e Autenticação

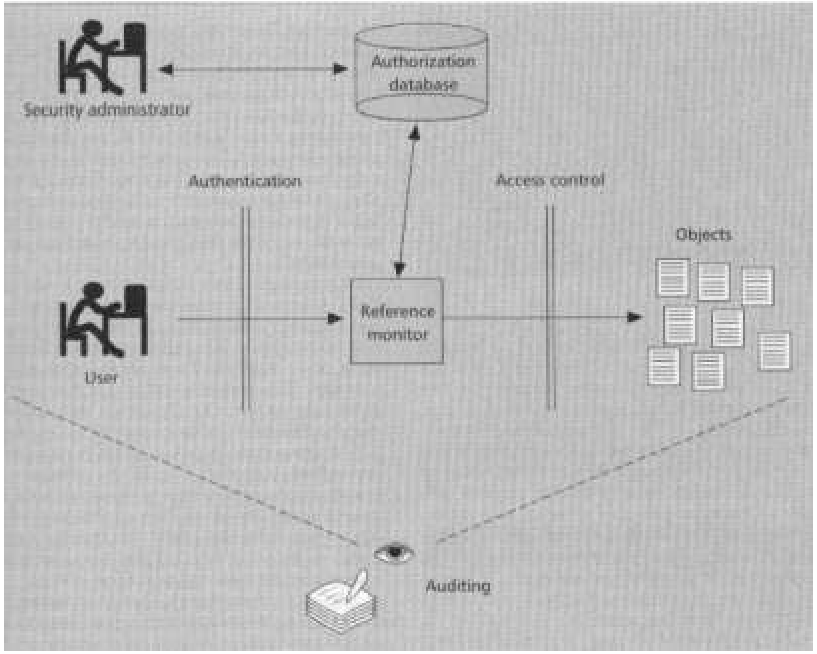
Henrique Santos (hsantos@dsi.uminho.pt)
Dpt. Sistemas de Informação
Ext. 510302

Summary

- Access Control
  - Models and Protocols
- Authentication
  - Password authentication
  - IP address based authentication
  - Cryptography based authentication
- User authentication
  - Passwords in detail
  - Tokens in detail
  - Biometrics in detail

Access Control

- Includes: Authentication, Authorization and Accounting/Auditing (AAA); but most of the times they are not fully implemented!



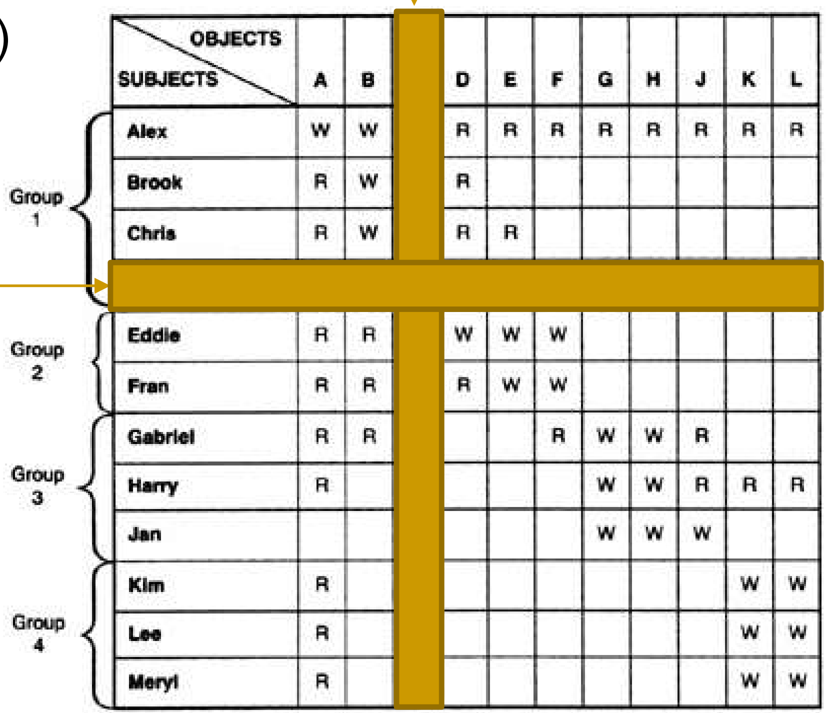
Access Control

- To control the access conditions of a ***subject*** to an ***object***, in particularly what the first can do (authorization) – Read, Write, Execute…
- Two implementation models
  - Based on Access Matrix – Process typically managed by the Operating System
  - Based on the attribution of capabilities – Process typically managed by a central server

Access Control

Access Matrix implementation models

- **ACL (Access Control List)**
- CL (Capability List)

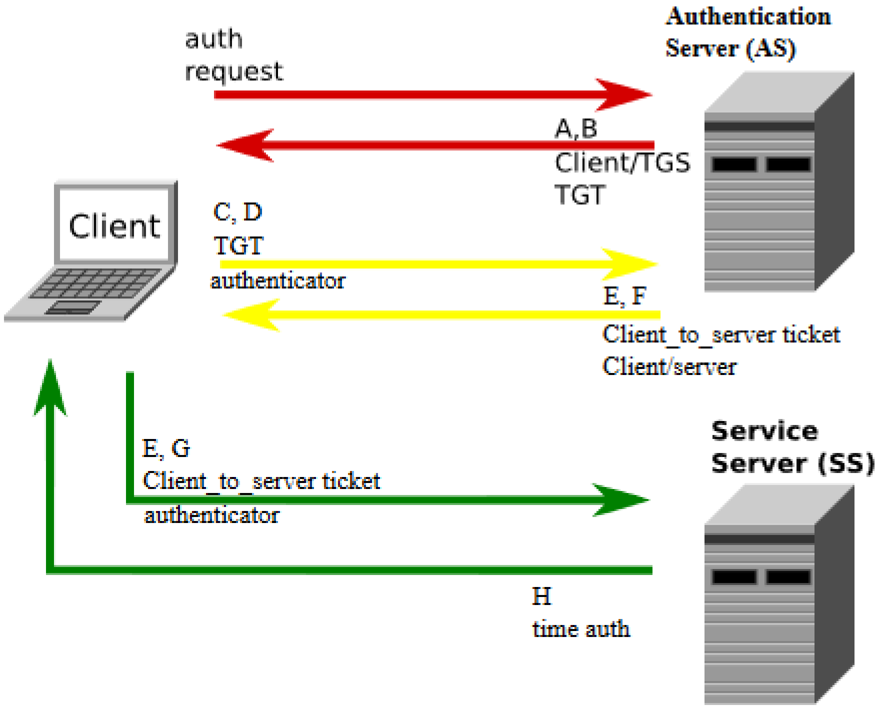


	OBJECTS / SUBJECTS	A	B		D	E	F	G	H	J	K	L
Group 1	Alex	W	W		R	R	R	R	R	R	R	R
	Brook	R	W		R							
	Chris	R	W		R	R						
Group 2	Eddie	R	R		W	W	W					
	Fran	R	R		R	W	W					
Group 3	Gabriel	R	R				R	W	W	R		
	Harry	R						W	W	R	R	R
	Jan							W	W	W		
Group 4	Kim	R									W	W
	Lee	R									W	W
	Meryl	R									W	W

Notes:
R Read
W Write and read

Access Control

Attribution of Capabilities – ticket based (Kerberos, Active Directory,…)



Access Control

- Main policies
  - Discretionary Access Control (DAC)
    - The object's access policy is defined by the owner
  - Mandatory Access Control (MAC)
    - The object's access policy is defined by the system (rigid, typically used in multi-level system, where the subjects and objects are masked by sensitivity security labels) – *need-to-know principle*
  - Role based Access Control (RBAC)
    - Like in MAC, the access policy is defined by the system. But instead of having permissions associated with subject's security levels, permissions are associated with subject's roles in the system.
  - Attribute based Access Control (ABAC)
    - Based on user's attributes (e.g., "older then 18"); XACML (eXtensible Access Control Markup Language) is a web standard since January, 2013.

SANDHU, R.S. 1993. LATTICE-BASED ACCESS-CONTROL MODELS. Computer 26, 9-19.

Sandhu, R.S., Coyne, E.J., Feinstein, H.L. and Youman, C.E. Role based access control models. Computer, 29 (2). 38-&.

AC Security Models

- Various types of formal specification models:
  - Confidentiality policy oriented (**Bell-LaPadula**), or integrity policy oriented (Biba, Clark-Wilson)
  - Models for static policies (**Bell-LaPadula**); vs. models that consider dynamic access rights (Chinese Wall)
  - Models can be informal (Clark-Wilson), semi-formal, or formal (**Bell-LaPadula**, **Harrison-Ruzzo-Ullman**).

Bell-LaPadula (BLP)

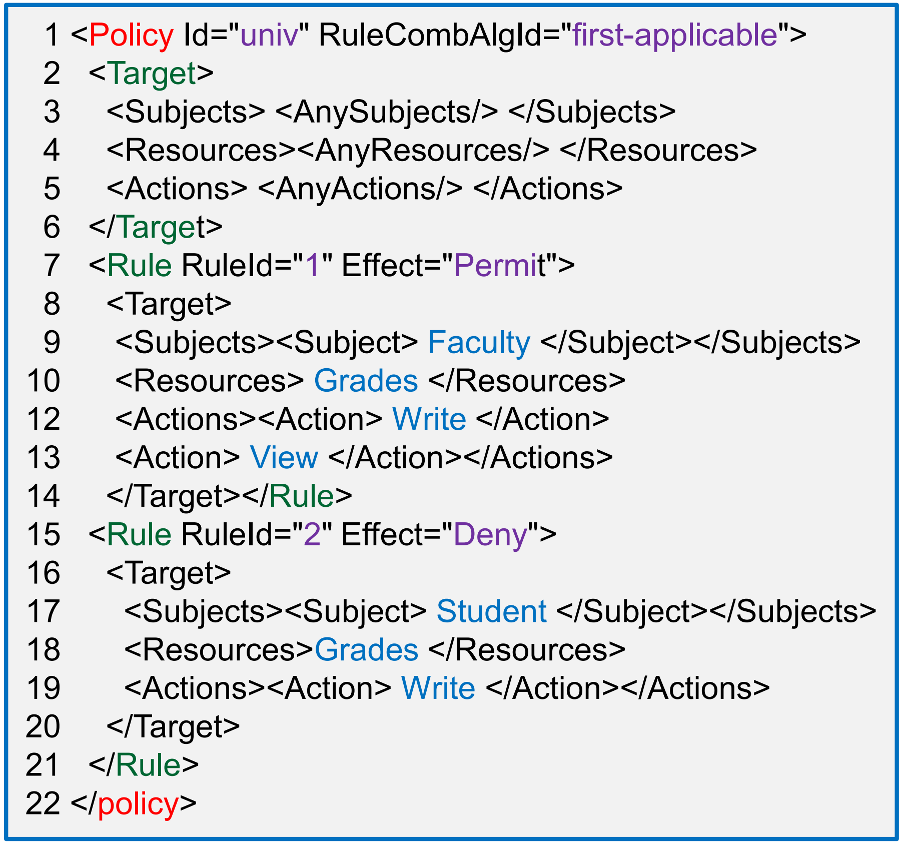
- Basis of several standards, including DoD’s Trusted Computer System Evaluation Criteria (TCSEC or “Orange Book”).
- It models confidentiality aspects of multi-user systems, e.g. in operating systems; combines aspects of DAC and MAC:
  - Access permissions are defined both through ACLs and through security levels
  - Multi-level security (MLS): mandatory policies prevent information flowing downwards from a high security level to a low-level one – **sanitization** operation required for practical implementations!
  - BLP is a static model: security levels (labels) never change.

Bell-LaPadula (BLP)

- BLP is a formal state transition model for computer security policies; it defines “secure states” and transitions, which preserve security.
  - The static nature is its main limitation… no policy for the creation and deletion of subjects and objects, or to change rights.
- The **Harrison-Ruzzo-Ullman model** defines (a limited set of) authorization procedures and objects with and without restrictions. Very complex but more close to OSs’ characteristics
  - http://en.wikibooks.org/wiki/Security_Architecture_and_Design/Security_Models

AC security policy specification

- XACML (eXtensible Access Control Markup Language)
  - Platform-independent
  - Proposed by OASIS
  - Rule-based, but several “profiles” have been proposed, i.e., model-base for RBAC
  - Example:



```
<Policy Id="univ" RuleCombAlgId="first-applicable">
  <Target>
    <Subjects> <AnySubjects/> </Subjects>
    <Resources><AnyResources/> </Resources>
    <Actions> <AnyActions/> </Actions>
  </Target>
  <Rule RuleId="1" Effect="Permit">
    <Target>
     <Subjects><Subject> Faculty </Subject></Subjects>
     <Resources> Grades </Resources>
     <Actions><Action> Write </Action>
     <Action> View </Action></Actions>
    </Target></Rule>
  <Rule RuleId="2" Effect="Deny">
    <Target>
      <Subjects><Subject> Student </Subject></Subjects>
      <Resources>Grades </Resources>
      <Actions><Action> Write </Action></Actions>
    </Target>
  </Rule>
</policy>
```

Exercise

- In an university context, construct the lattice of security labels for the security levels P (**public**), C (**confidential**) and SC (**strictly confidential**) – P < C < SC – and categories **AS** (Academic Services) and **ScS** (Scientific Services)
- Assuming:
  - the fundamental BLP model properties
  - teachers are classified at level (*label*) **(C, {AS, ScS})**
  - the usual model implementation (multilevel) on a computer system

  ascertain if it is possible to prevent a student classified as **(C, {AS})** cheating with a teacher.
- Elaborate about a **possible automatic deployment process** of such a model in a typical CIT infrastructure

Notes: you are required to understand formal aspects of BLP model; it may help to get familiar with SELinux and the Linux *acl command family.

Additional information:

- Sandhu, Ravi S. "Lattice-based access control models." *Computer* 26.11 (1993): 9-19.
- http://www.cs.cornell.edu/courses/cs5430/2011sp/NL.accessControl.html
- http://www.cs.unc.edu/~dewan/242/f96/notes/prot/node1.html

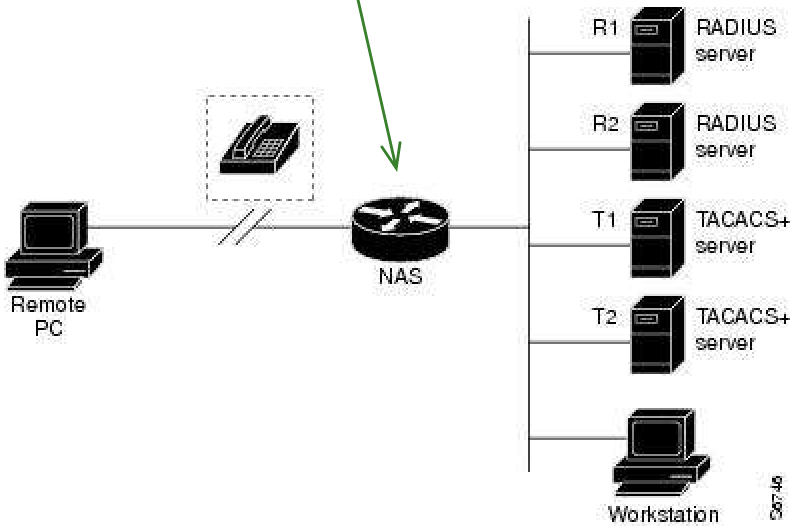
Access Control

- Widespread protocols
  - RADIUS – Remote Authentication Dial In User Service: is an AAA protocol (**application level**) particularly suitable to control access to network resources; combines authentication and authorization; based on UDP; cyphers only user password
    http://en.wikipedia.org/wiki/RADIUS
  - TACACS+ - Terminal Access Controller Access-Control System Plus: very similar to the previous; separates the operations of authentication and authorization operations; uses TCP; focus on device administration; cyphers all authentication process
    http://tools.ietf.org/html/draft-grant-tacacs-02
  - Kerberos – developed at MIT; a secret-key network authentication protocol; based on the concept of a centralized system for key distribution and user authentication; limited auditing capability
    https://web.mit.edu/kerberos/

Access Control

- Typical AAA Network Configuration
  - Multiple Security Servers (possible answers: FAIL; PASS; ERROR)
  - If anyone returns “FAIL” access is denied

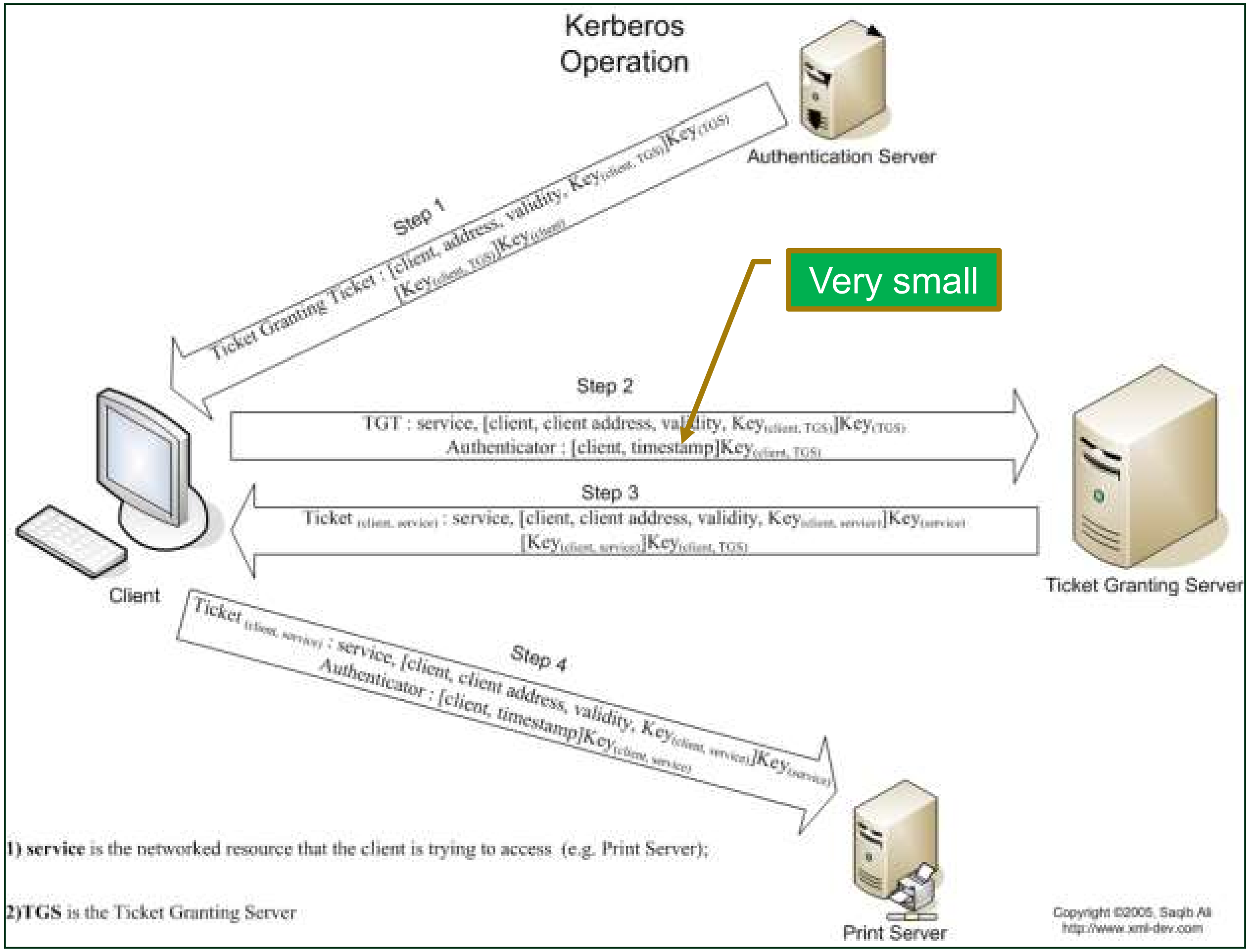
e.g.:
OpenNAC
PacketFence
FreeNAC



http://www.cisco.com/c/en/us/td/docs/ios/12_2/security/configuration/guide/fsecur_c/scfaaa.html

KERBEROS

- protocol summary



Kerberos Operation

Step 1

Ticket Granting Ticket : [client, address, validity, $Key_{(client, TGS)}$]$Key_{(TGS)}$ [$Key_{(client, TGS)}$]$Key_{(client)}$

Step 2

TGT : service, [client, client address, validity, $Key_{(client, TGS)}$]$Key_{(TGS)}$
Authenticator : [client, timestamp]$Key_{(client, TGS)}$

Step 3

Ticket $_{(client, service)}$: service, [client, client address, validity, $Key_{(client, service)}$]$Key_{(service)}$ [$Key_{(client, service)}$]$Key_{(client, TGS)}$

Step 4

Ticket $_{(client, service)}$: service, [client, client address, validity, $Key_{(client, service)}$]$Key_{(service)}$
Authenticator : [client, timestamp]$Key_{(client, service)}$

Very small

1) **service** is the networked resource that the client is trying to access (e.g. Print Server);

2)**TGS** is the Ticket Granting Server

Copyright ©2005, Saqib Ali
http://www.xml-dev.com

KERBEROS - summary

- Passwords are not communicated
- All communications are encrypted
- Limited validation period
- Timestamps to prevent attacks by reuse
- Mutual authentication

- Evolution: version 4 (using DES) has established itself as a standard for the Internet. Version 5 (not restricted to DES) addressed some security issues

KERBEROS - summary

- But…
- Single point of failure: **Ticket Granting Server**
- Authentication between the TGS and all servers is done through a secret key! ... Scalability is limited
- Timing issues can become critical (synchronous clocks and tickets' validity time)
- A workstation can save the password and use it later!
- Messages exchanged with the AS can be captured and subjected to a password attack (brute force or dictionary)

Authentication

- Authentication
  - Subject’s identity verification process, **with a certain degree of confidence** (the subject can be a human or a machine).
  - Two typical cases:
    - A computer requires access to another shared computer.
    - A user requires access to an workstation

Password Authentication

- A password is a secret shared between entities that require some level of confidentiality
    - Man-Machine
        - It demands memorization by the Man ☹
        - Typically managed and controlled by the Man ☹
            - Weak passwords (possibly found in dictionaries)
            - The same password is used in several relations (exposition)

Password Authentication

- Passwords (cont)
  - Machine-Machine
    - It may be a much more elaborated secret ☺
    - But, unfortunately, often based on the mechanisms used in the previous case ☹
  - Transmission channel should also be considered (encryption)

Password Authentication

- Password storing
  - How do servers know passwords?
    - Each one has a copy;
    - There is a central repository where every computer can look for passwords; and
    - There is a central server that does the Authentication and informs the others

- Difficult to maintain
- A compromised server does not compromise the others

- Easy to maintain
- Single point of failure...
- But allows to focus the security efforts

Password Authentication

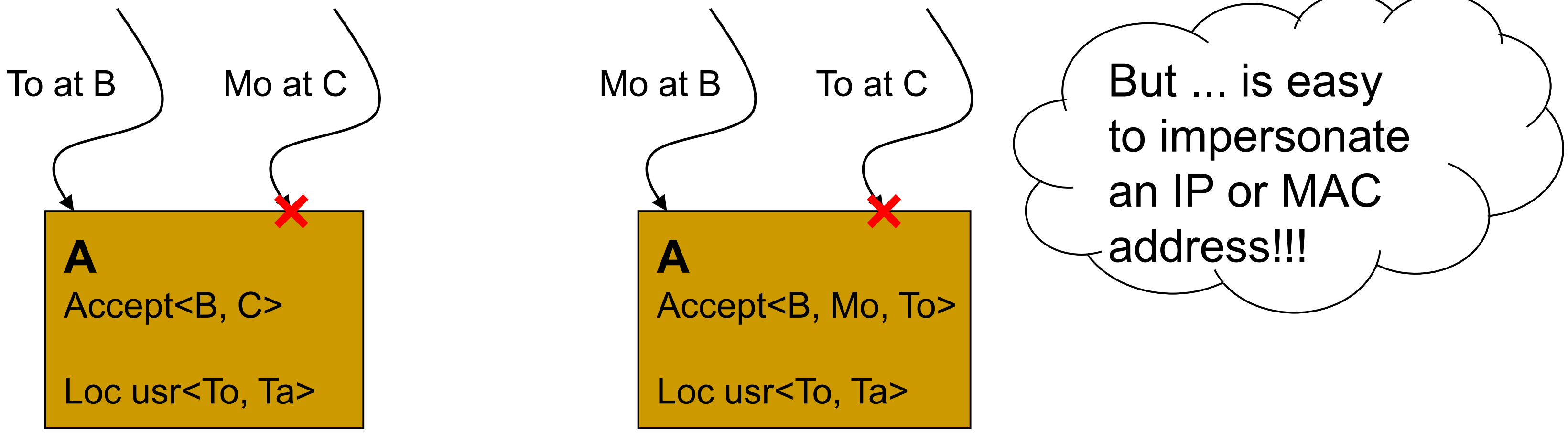
- Password storing
  - Encrypted file
    - If the encryption key is known all passwords are compromised
  - Passwords protected individually (store a password hash instead of the password itself – UNIX and VMS)
    - Possible disclosure of one does not affect the others...
  - Mixed solutions

Password Authentication

- Password storing (cont)
  - A Directory Service is very common
    - Active Directory (MS); NIS – Network Information Service (SUN)
    - …
    - Frequently the server does not authenticate itself (vulnerability)

Authentication based on net Address

- Subject’s **identity can be inferred** from the network address 😐
  - Each machine has a list of allowed relationships. Examples:



To at B

Mo at C

A

Accept<B, C>

Loc usr<To, Ta>

Mo at B

To at C

A

Accept<B, Mo, To>

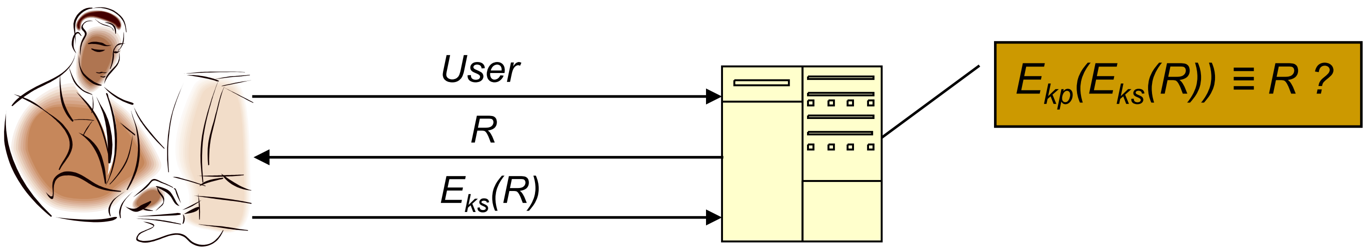
Loc usr<To, Ta>

But ... is easy to impersonate an IP or MAC address!!!

- Forces to have the same login on all machines
- Management of equivalent logins is very complex

Cryptographic authentication

- Authentication protocols based on encryption
  - Public key algorithms guarantee authentication without transmitting the key or any password. Basic idea:



  - Symmetric encryption techniques are also possible, but are more complex ... **Kerberos** is an example
  - **Efficiency relies on the key generation technique**
    - Simpler when limited to machines
    - For users, it is desirable to deduce the key from a password

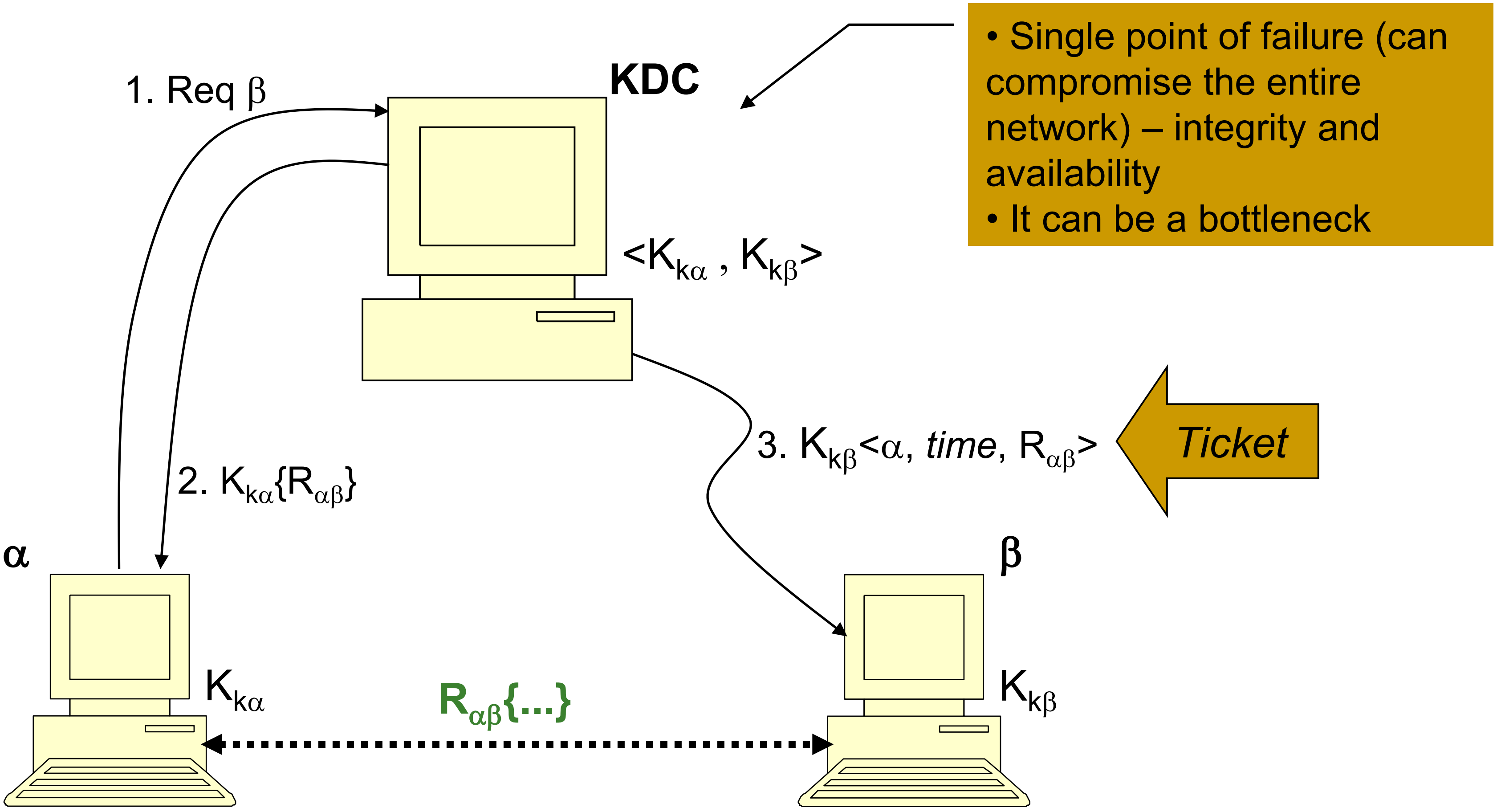
Cryptographic authentication

- The password can be used to derive the key
  - A Hash function applied to the password
  - Converting the password into a cryptographic key
    - Symmetric key is easy…
    - A key pair is more complicated, but one possible solution is to use the password as a seed to a random number generator ... computationally heavy
    - Use the password to decrypt the private key, obtained from a directory service (for example)
    - The actual password is the main vulnerability (and under responsibility of the user ☹)

Cryptographic authentication

- Authentication between machines on the network, using symmetric keys
  - N machines ⇒ N – 1 keys stored in each node!
  - What happen when a new node is added? (scalability)
- How can we distribute keys?
  - Using a Key Distribution Center (KDC), which shares a secret key with every nodes.

Cryptographic authentication



- Single point of failure (can compromise the entire network) – integrity and availability
- It can be a bottleneck

KDC

$<K_{k\alpha}, K_{k\beta}>$

1. Req β

2. $K_{k\alpha}\{R_{\alpha\beta}\}$

3. $K_{k\beta}<\alpha, \textit{time}, R_{\alpha\beta}>$

Ticket

α

$K_{k\alpha}$

β

$K_{k\beta}$

$R_{\alpha\beta}\{...\}$

Cryptographic authentication

- It is more efficient with public key cryptography
  - Each node has its private key
  - All public keys are stored centrally
  - How to ensure the association of a public key with an entity?
    - Digital certificates signed by a CA (**Certification Authority**)
    - Each node has the CA's public key
  - But there remains a question...
    Who is using the certificate? Is it the owner? ☹

User Authentication

- Something the user knows (Knowledge-based)
  - Passwords

- Something the user has (Object-based)
  - *Tokens*

- Something the user is (ID-based)
  - Biometrics

User authentication methods

- User level acceptance (Jones's study)
  - Keyword is the best known mechanism, followed by some biometrics, and finally, tokens
  - Preferences:
    - Computer access – **password**
    - Financial transactions – **passwords and biometrics**
    - Health activities – **Biometrics**
    - Physical access – Tokens

User authentication methods

- User level acceptance (cont)
  - Biometrics in financial transactions
    - Fingerprint; digital signature analysis; hand geometry
  - Perception of security
    - Biometrics (iris; fingerprint; hand geometry; voice and face recognition;…), followed by passwords and, at last, tokens
  - Impact on privacy
    - There are no key differences (biometrics; keyword; tokens)

User Authentication

<table>
<tr><td>Usually referred to by:</td><td><b>Password; Secret</b></td><td><b><i>Token;</i> Card</b></td><td><b>Biometric</b></td></tr>
<tr><td>Authentication based on:</td><td>Secrecy or obscurity</td><td>possession</td><td>Individualization and personalization</td></tr>
<tr><td>Security assumption:</td><td>It is never revealed</td><td>It is never lost</td><td>Unable to duplicate</td></tr>
<tr><td>Example (digital):</td><td>Computer access password</td><td>Card access garage</td><td>Fingerprint</td></tr>
<tr><td>Security limitations:</td><td>Less safe with use; memorization</td><td>Compromised if it is lost</td><td>Very hard to replace</td></tr>
<tr><td rowspan="4">Combinations (multifactor)</td><td colspan="2">Two-factor authentication</td><td></td></tr>
<tr><td></td><td colspan="2">Two-factor authentication</td></tr>
<tr><td>Two-factor</td><td></td><td>authentication</td></tr>
<tr><td colspan="3">Three-factor authentication</td></tr>
</table>

User authentication methods

- Passwords in detail
- Tokens in detail
- Biometrics in detail

Passwords in detail

- Vulnerabilities
  - Could be guessed
  - Could be forgotten
  - Could be shared
  - Could be written down and subsequently lost or stolen
- Attack origin
  - On-line – trying to avoid
    - Limit the number of attempts
    - Suspect of larger number of attempts (**auditing**)
  - Off-line – trying to avoid
    - Protect stored passwords
    - Promote the use of strong passwords
    - One-Time passwords (or challenge-response mechanisms)

Passwords in detail

- Attack methods
  - Guessing (pre-knowledge and common passwords); dictionary; brute force
    - **Password size is critical**
  - Even strong passwords are exposed
    - Key loggers
    - Phishing attacks
    - Shoulder surfing attacks
  - Eavesdropping: direct observation or communication sniffing
    - One-time passwords can help
    - Never communicate passwords in clear text

Passwords in detail

- Attack methods (cont)
  - Careless users
    - Registration on paper in a public place
    - Using the same password on multiple systems
    - Let be deceived by Trojans and Phishing
    - Leave terminals logged
    - Shoulder surfing attacks

  Most of these risks are minimized through proper management (password creation, renewal, etc.)

Tokens in detail

- Contains authentication information
- Can implement strong passwords
- **Can be stolen or lost**, and therefore require an authentication mechanism for the user (typically a PIN – Personal Id Number)
- Several types:

Biometrics

- More than a century has passed since Alphonse Bertillon devised and "industrialized" an idea to identify criminals using data from the body.
- In 1893 the United Kingdom Ministry of Internal Affairs “assumes” that no two individuals have the same fingerprint.
- The first AFIS (Automatic Fingerprint Identification System) appeared in 1960.
- In recent decades many techniques have emerged. With the help of Hollywood (CSI) it emerged the idea that biometry has a set of very mature techniques!
    - In 2004, from a competition on AFIS it was revealed that the best techniques generated 2% false negative!

Biometric types

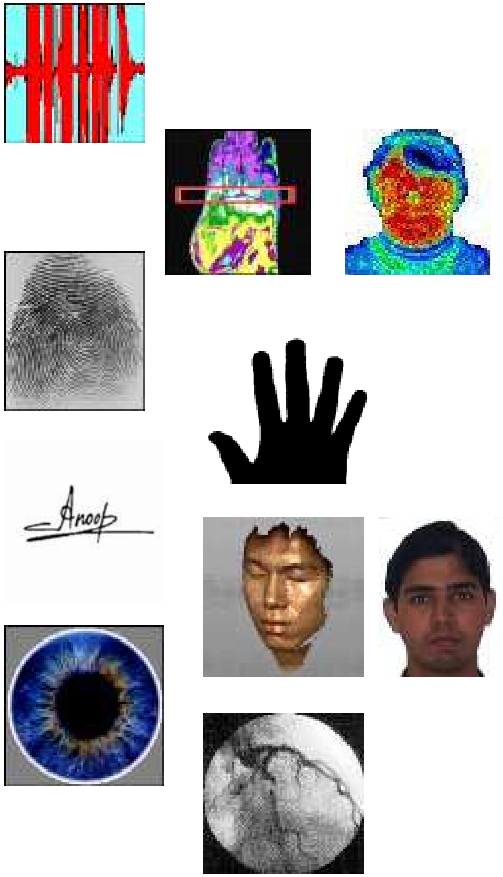
- Which biological characteristics can be used?
  - Fundamental Properties:
    - Universality
    - **Distinctiveness** (uniqueness)
    - **Permanence** (immovability)
    - Collectability

Factors:
- Behavioural
- Genetics
- Random

  - Other requirements
    - Performance (accuracy, resources, etc.)
    - Acceptability
    - Circumvention (resistance to direct attacks)

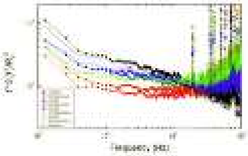
Biometric types

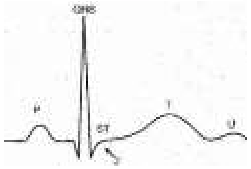
- Well established
  - Voice
  - Infrared thermograms: facial analysis and hand’s veins pattern
  - Fingerprint
  - Hand geometry
  - Signature
  - Face
  - Iris
  - Retina

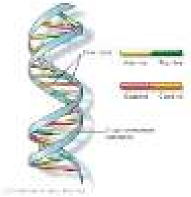


Biometric types

- Under research
  - Keystrokes dynamics
  - Gait
  - Odor
  - Ear
  - Electrocardiogram
  - DNA
  - Multidimensional

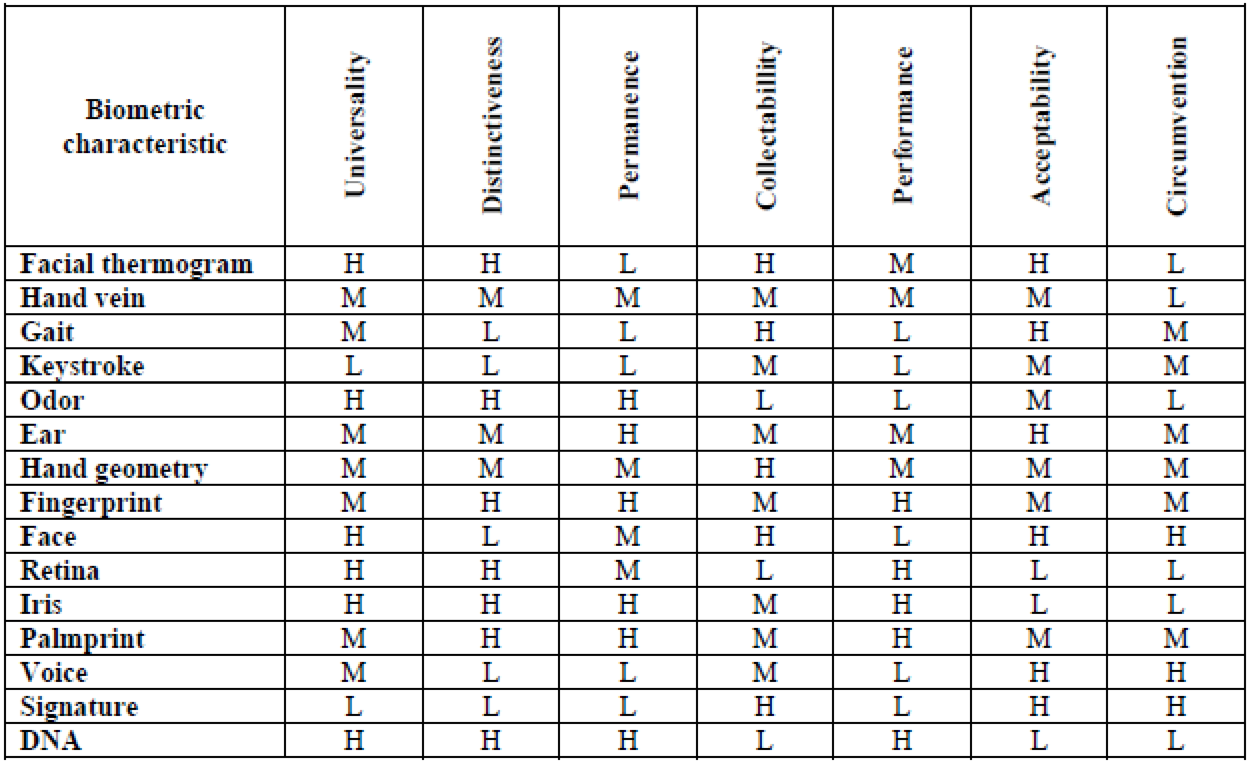






Biometric types

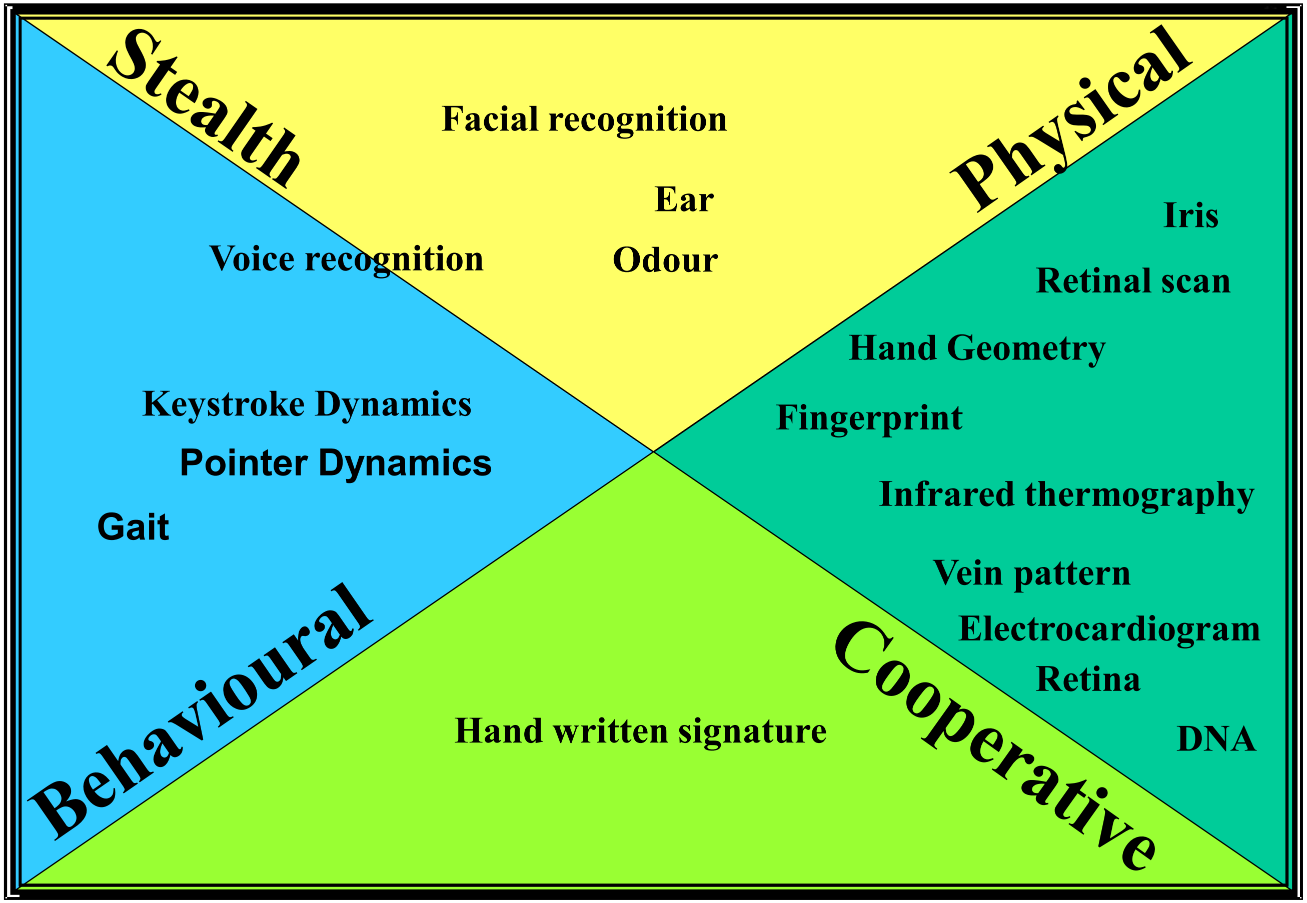
- Which are the best characteristics?



Biometric characteristic	Universality	Distinctiveness	Permanence	Collectability	Performance	Acceptability	Circumvention
Facial thermogram	H	H	L	H	M	H	L
Hand vein	M	M	M	M	M	M	L
Gait	M	L	L	H	L	H	M
Keystroke	L	L	L	M	L	M	M
Odor	H	H	H	L	L	M	L
Ear	M	M	H	M	M	H	M
Hand geometry	M	M	M	H	M	M	M
Fingerprint	M	H	H	M	H	M	M
Face	H	L	M	H	L	H	H
Retina	H	H	M	L	H	L	L
Iris	H	H	H	M	H	L	L
Palmprint	M	H	H	M	H	M	M
Voice	M	L	L	M	L	H	H
Signature	L	L	L	H	L	H	H
DNA	H	H	H	L	H	L	L

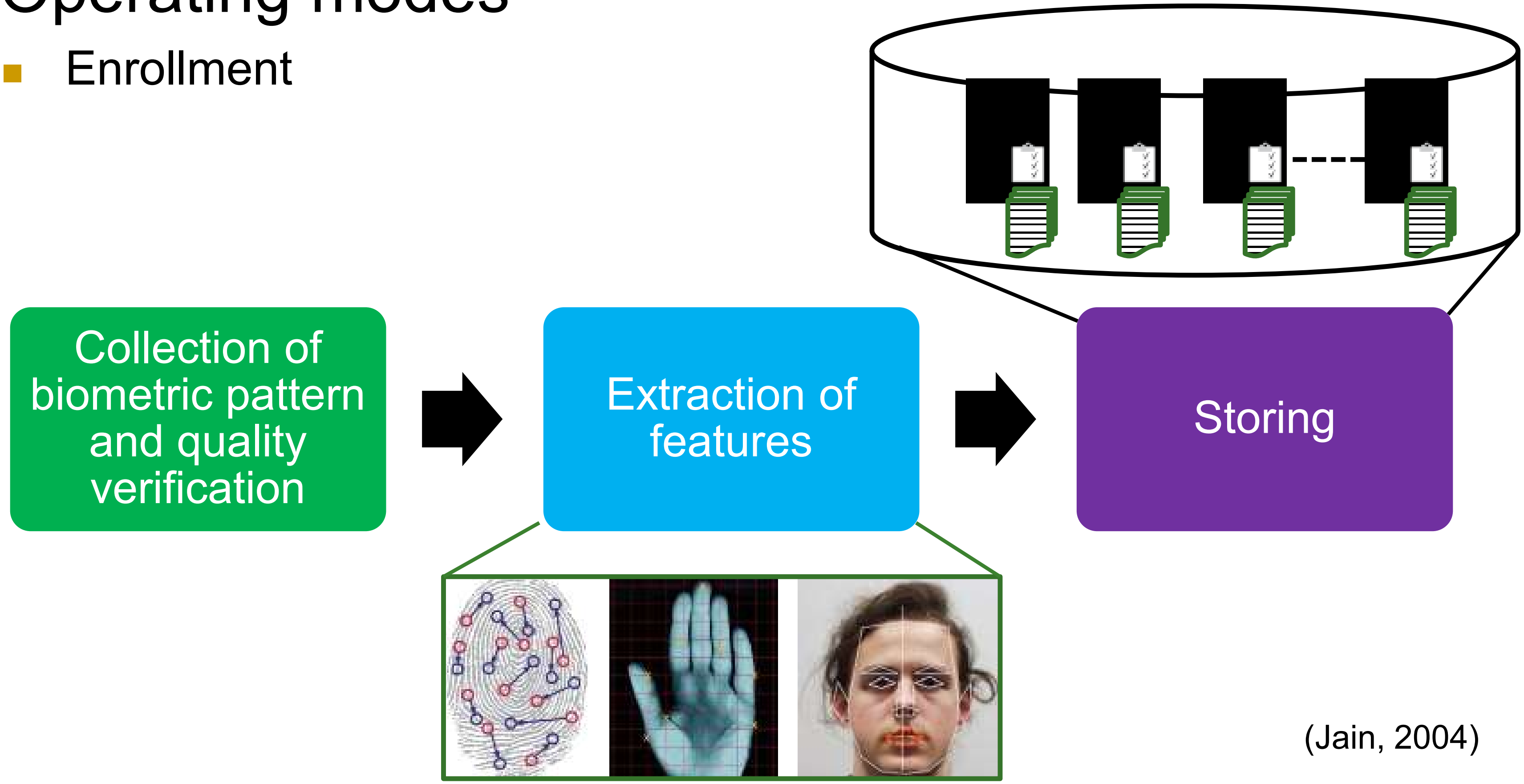
(Delac, 2004)

Biometrics - taxonomy



Biometric Systems

- Operating modes
  - Enrollment



Collection of biometric pattern and quality verification

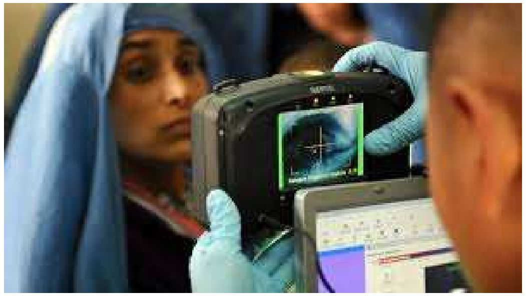
Extraction of features

Storing

(Jain, 2004)

Biometric Systems

- Sensor
  - Collect raw data, eventually with quality verification
  - Fingerprint, face and iris are the most well known
  - Some signal processing techniques (filtering) and image processing techniques (specialy when using images or video)

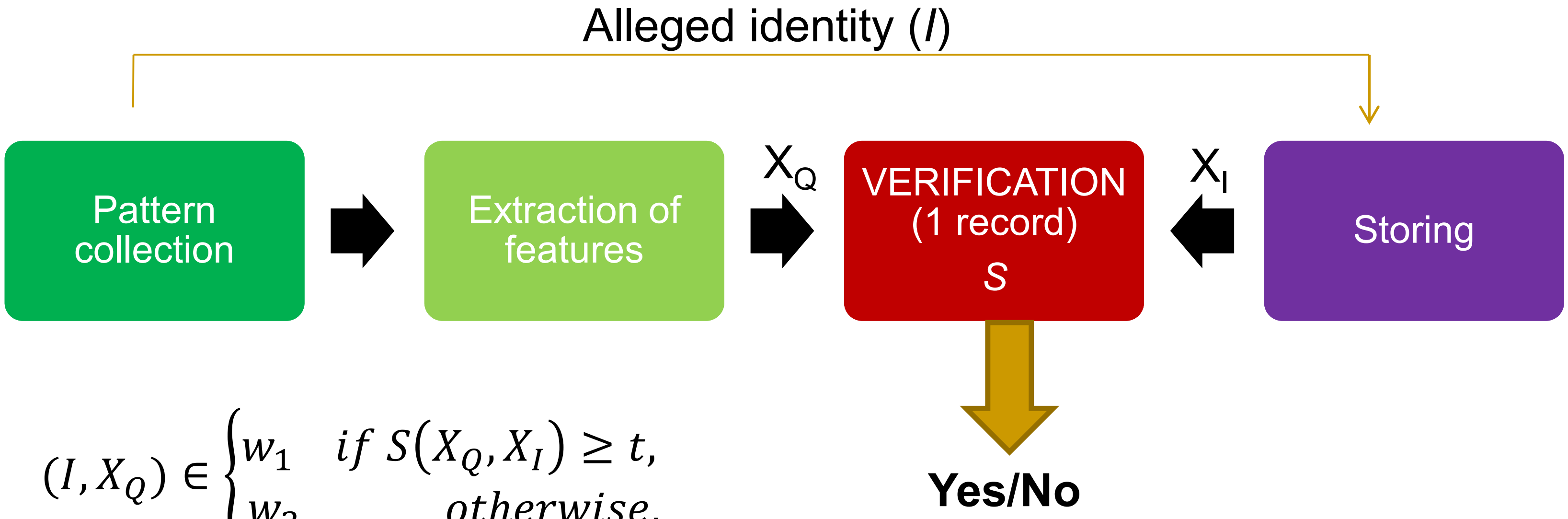


Biometric Systems

- Features extraction
  - Pattern recognition problem
  - Machine learning techniques used with some success:
    - Principal Component Analysis – *Eigenfaces*
    - Gabor Filters
    - Linear Discriminant Analysis – LDA
    - Naive Bayes Classifier
    - Rough Sets
    - Neural Networks
    - Support Vector Machines
    - …
  - Supervised… training is critical

Biometric Systems

- Operating modes
  - Verification (positive recognition)
    - The individual is who he/she claims to be? (e.g., system authentication)



$$(I, X_Q) \in \begin{cases} w_1 & if\ S(X_Q, X_I) \geq t, \\ w_2 & otherwise, \end{cases}$$

Biometric Systems

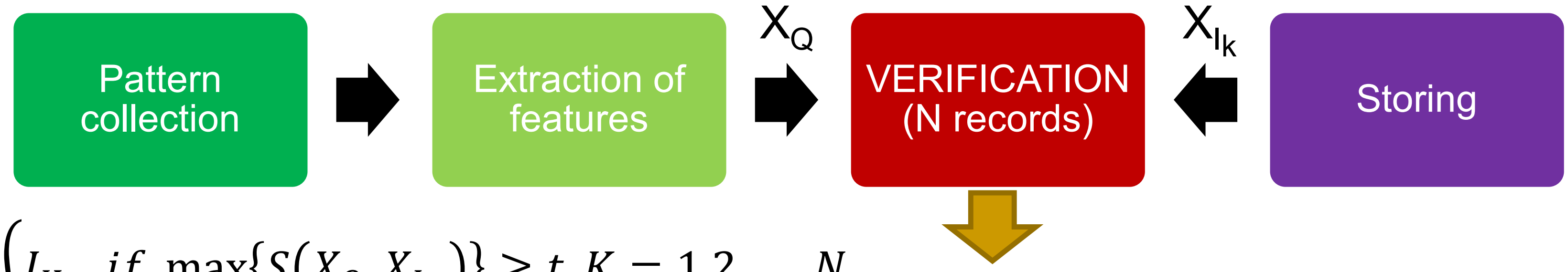
- *S*: Similarity function (produce a *matching score*), typically:
  - Euclidian distance $S = \sqrt{\sum_{i=1}^{n}\left(X_{Qi} - X_{Ii}\right)^2}$
  - Mahalanobis distance $S = \sqrt{\left(\overrightarrow{X_Q} - \overrightarrow{X_I}\right)^T S^{-1}\left(\overrightarrow{X_Q} - \overrightarrow{X_I}\right)}$

    where S^{-1} is the covariance invert matrix, or precision matrix
  - Manhattan distance (taxicab metric, or rectilinear distance) $s = \|p - q\| = \sum_{i=1}^{n}|p_i - q_i|$

    where $p = (p_1, p_2, \ldots, p_n)\ and\ q = (q_1, q_2, \ldots, q_n)$ are vectors
  - Camberra distance (variant of taxicab metric) $s = d(p, q) = \sum_{i=1}^{n}\frac{|p_i - q_i|}{|p_i| + |q_i|}$
  - Hamming distance
- Effect of variation (random) of X_Q, or even X_I
- *t*: it is a pre-defined *threshold*
- **In any case the model demands for large studies of the target population**

Biometric Systems

- Operating modes
  - Identification (negative recognition) – only possible with biometrics
    - From a biometric pattern, is the individual already registered? (e.g., driving license request)
  - Detection (particular case of identification)
    - This biometric pattern belongs to an individual included on a "wanted" list? (e.g., airport security, or e-Passport)

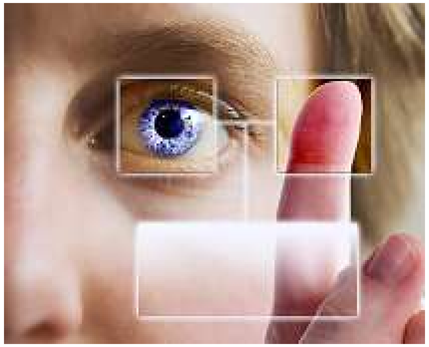


$$(X_Q) \in \begin{cases} I_K & if \ \max_K\{S(X_Q, X_{I_K})\} \geq t, K = 1,2, \dots, N, \\ I_{N+1} & otherwise, \end{cases}$$

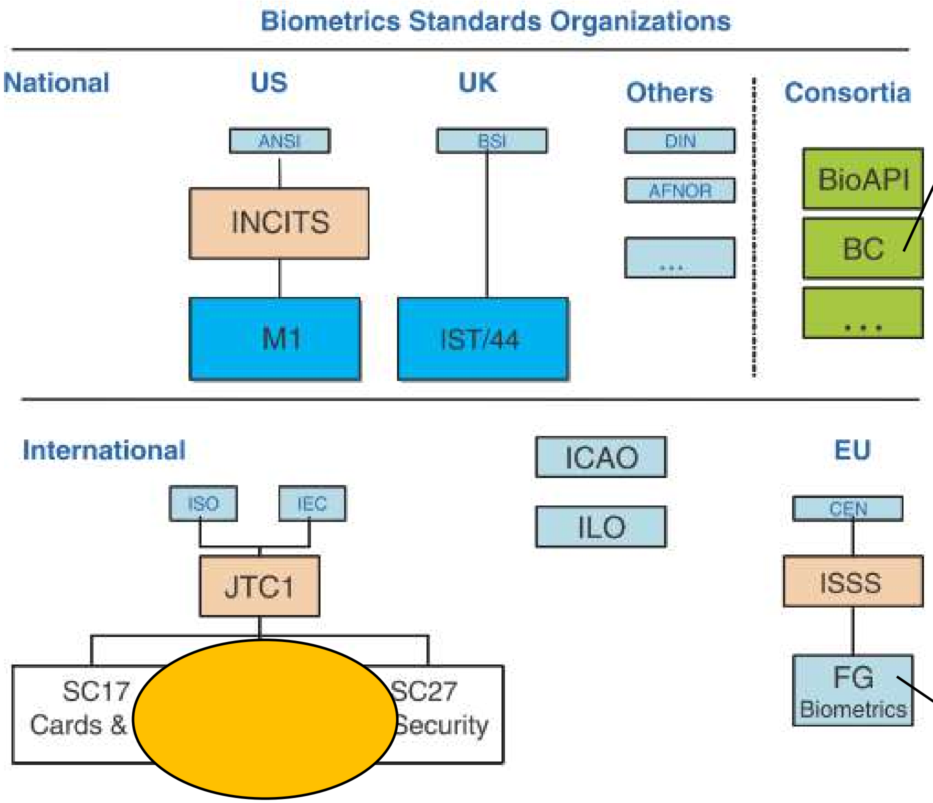
Identified / Unidentified

Biometric Systems - storage

- Biometric patterns with…
  Quality indicators
  Context (sensors, algorithms, etc.)
  Identity
  Raw data (for study and evaluation purposes)
  …
- Available data bases:
  - CASIA / Biometrics Ideal Test (http://biometrics.idealtest.org/)
  - FERET among others, for face recognition: http://www.face-rec.org/databases/
  - Used within international competitions (http://www.nist.gov/biometrics-portal.cfm)
- Secure storage
  - Cryptography



Biometrics - Standardization



Biometric Consortium (NSA e NIST…)

Focus Group Biometrics, (now closed)

(Deravi 2008)

Biometrics - challenges

- Accuracy and evaluation
- Scalability
- Security
- Privacy

Accuracy and evaluation

- Types of evaluation
  - Technological
    - Needs a clean and normalized test data base; repeatable; algorithms evaluation
  - Operational
    - Real-time data; environment is not replicable; system performance evaluation
  - Scenario
    - Real data (reusable if the capture is controlled); complete system performance evaluation, using an application prototype and/or a simulated environment
- There are differences but the tools are the same
- More critical concerning Identification, but also relevant for Verification (Authentication)

(Gamassi, Lazzaroni et al. 2005)

Accuracy and evaluation

- Problem: discrete decision (accept/reject) based on probabilistic data, under the definition of a given *threshold*.

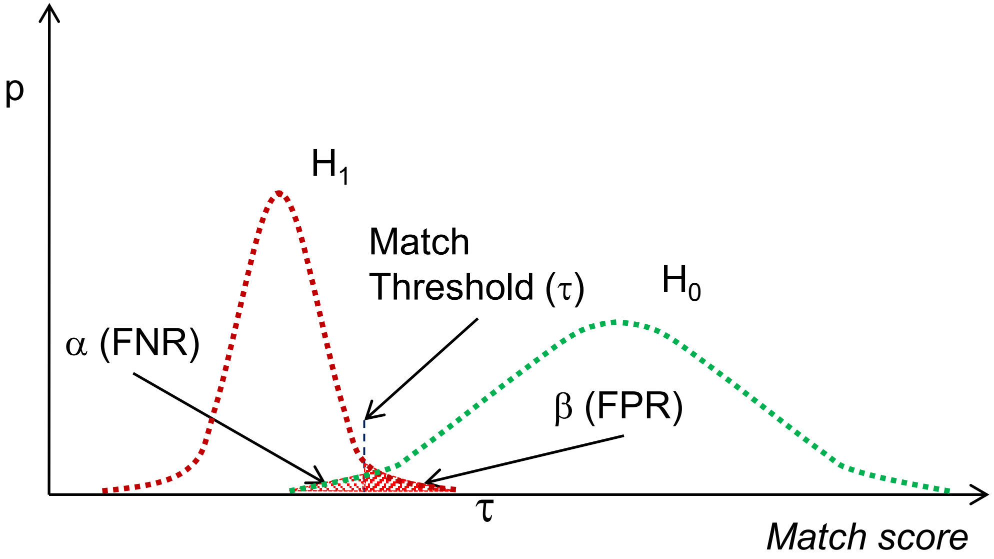
  “**What is the probability of the verification system make a wrong decision?**”

  Formulation: *Hypothesis testing*

  - *Null hypothesis* (H_0): the claimed identity is true (“genuine”)
  - *Alternative hypothesis* (H_1): the claimed identity is false (“impostor”)
  - *Test statistic*: typically a scalar value (*score*) that embraces all the (“noisy”) decision supporting information.
  - Result: not reject H_0; or reject H_0 in favor of H_1

Biometrics - verification

- Example of possible probabilistic density functions of similarity values for “genuine” (H_0 true) and “impostors” (H_1 true)
- Overlapped area is the source of decision errors – ***threshold* definition is critical**
  - **Type I errors** – when H_0 is true, but the decision is negative (FN or FR)
    The probability of a FN occurrence is given by α and denoted by FNR
  - **Type II errors** – when H_0 is false, but the decision is positive (FP or FA)
    The probability of a FP occurrence is given by β and denoted by FPR

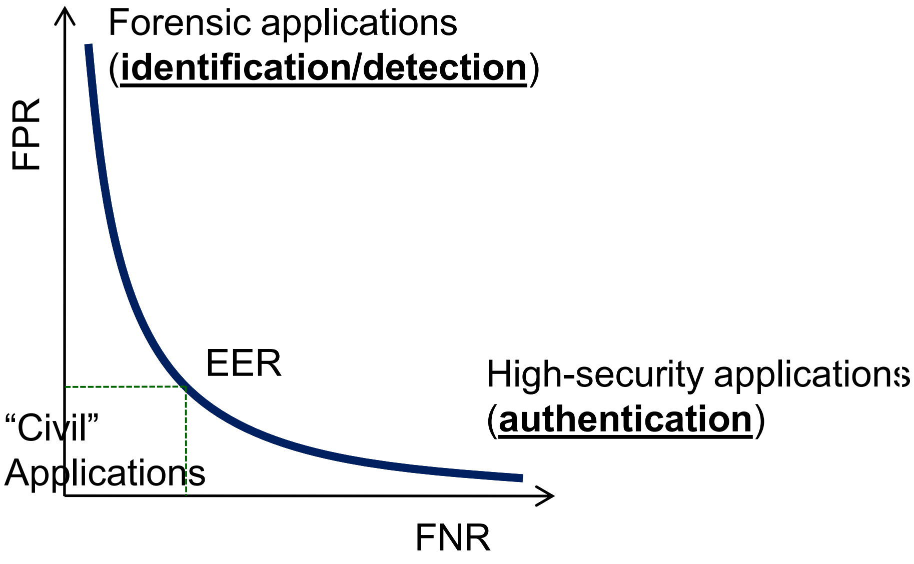


$$\beta = \int_{\tau}^{+\infty} f_{H1}(S)ds$$

$$\alpha = \int_{-\infty}^{\tau} f_{H0}(S)ds$$

Biometrics – verification (DET curves)

- FPR and FNR vary inversely depending on τ



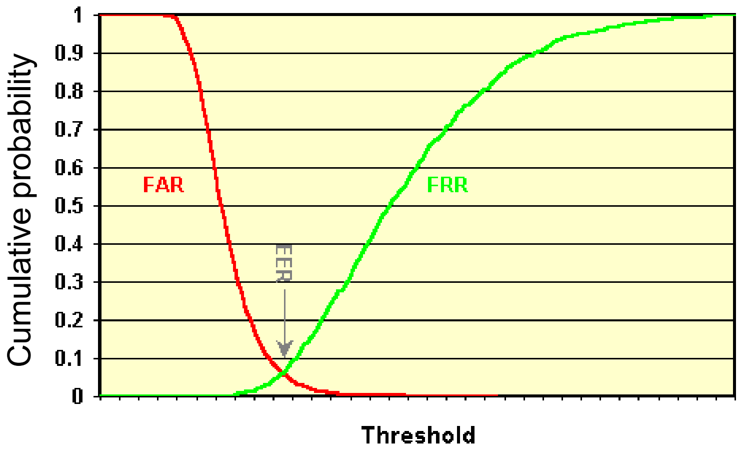
Detection Error Trade-off (DET) (Delac, 2004)

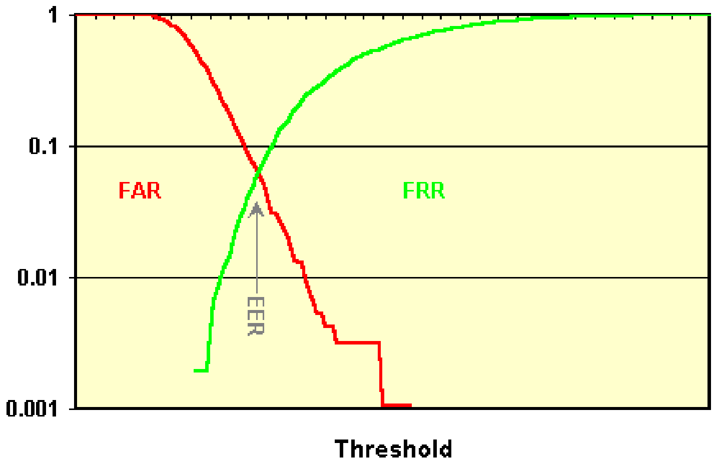
- EER – *Equal Error Rate* (resumes in a *simple* value, a possible performance indicator!)
  - But $\text{EER}_A < \text{EER}_B \nRightarrow$ A is better then B

Note: FPR and FNR are non-stationary statistic values!

Biometrics – verification (DET curves)

- Another way of representing DET curves (examples with linear and logarithmic scales)





Biometrics – verification (global evaluation)

- Other typical definitions in a decision binary system
  - TA – *hits*, or true positives
  - TR – true negatives, or correct rejections
  - FR – false rejections (*type I error*)
  - FA – false acceptations (*type II error*)

    M (total of legitimates) = TA + FR ⇔ TA = M - FR and

    NM (total of impostors or attacks) = TR + FA ⇔ TR = NM - FA
  - TAR = TA/M = 1 - FRR – sensibility
  - TRR = TR/NM = 1 - FAR – specificity
  - ACC = (TA + TR)/(M + NM) – precision

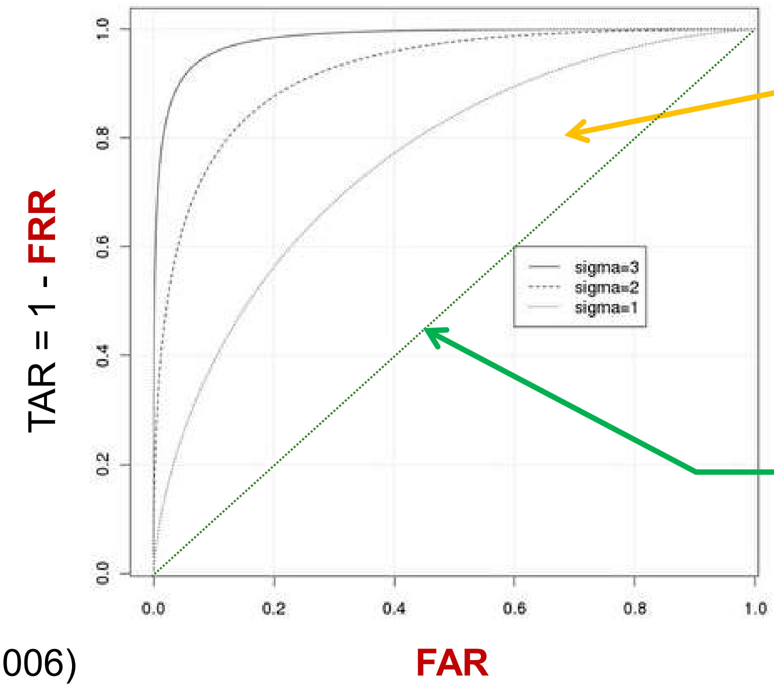
Confusion Table

TA	FR	→ Legitimates (M)
FA	TR	→ Impostors (NM)
↓ Accept	↓ Reject	

(Bewick, Cheek et al. 2004) e
(Ratha and Govindaraju 2008)

Biometrics –verification (ROC curves)

- The ROC curves (*Receiver Operating Characteristic*) are useful to relate FAR with FRR



Area under the ROC curve is an indicator of the global *performance*

Performance limit for any useful decision algorithm

(Fawcett 2006)

Biometrics – Individual evaluation

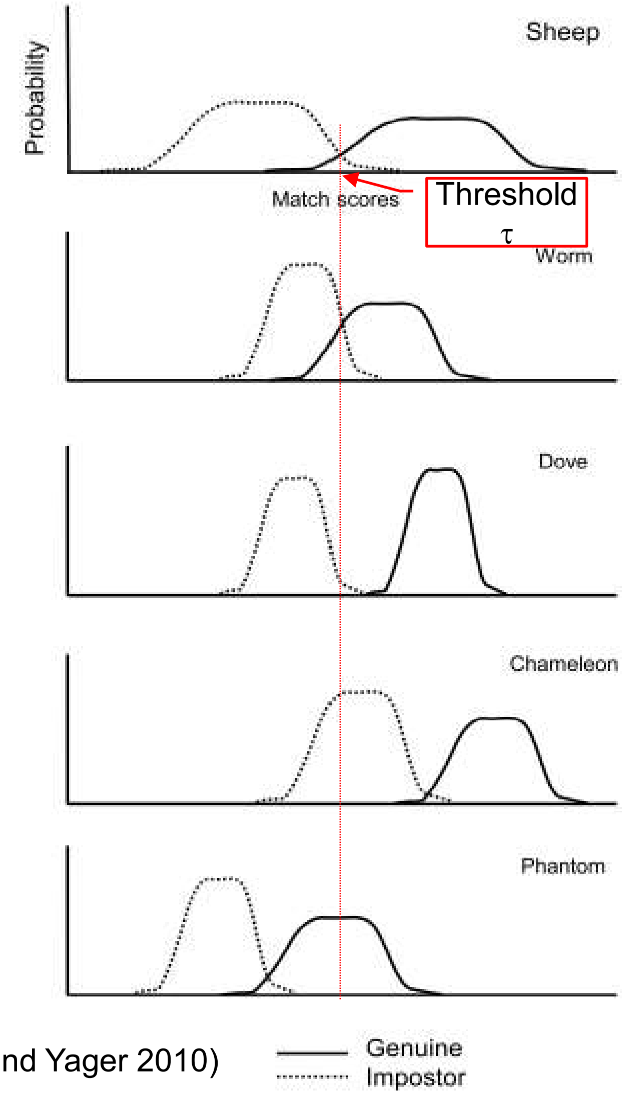
- Global analysis limitations (aggregated data)
- Individual factors affecting evaluation
  - Physiological
  - Behavioural
  - Interaction
- Individual analysis aiming threshold value τ
- This analysis conducted to the *Biometric Menagerie* (Yager, 2010)

Biometrics – Individual evaluation

- Classification based on **global evaluation**
  - *Sheep* – the most frequent (normal behaviour)
  - *Goats* – high FNM (low scores)
  - *Lambs* e *Wolves* – high FA (low scores as genuine; high scores as attacker)

Biometrics – Individual evaluation

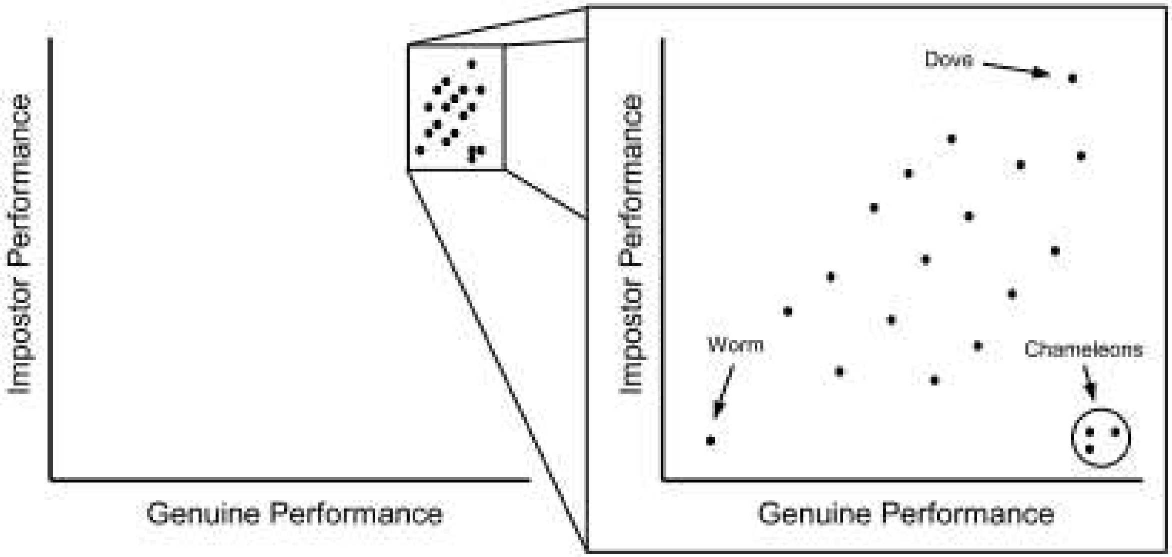
- Classification based on **individual distribution**
  - *Worms* – the worst distribution
  - *Doves* – (near) ideal distribution
  - *Chameleons* – easy impersonation against others
  - *Phantoms* – hardly authenticate



adapted from (Dunstone and Yager 2010)

Biometrics – Individual evaluation

- *Zoo Plot* (performance as genuine and against impostors); scale effect must be considered to identify groups



Biometric evaluation

Case study

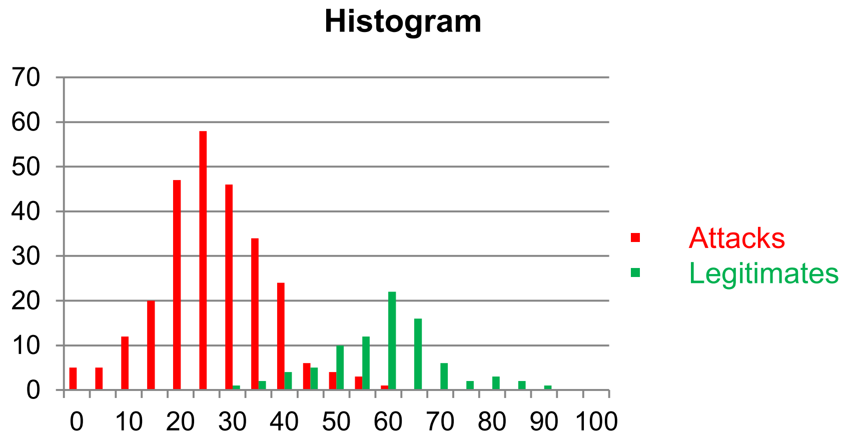
Biometrics – evaluation limitations

- Biometric performance
  - How to find a priori probability density functions? Not typical distributions that must be determined empirically. The gathering of samples is a key process:
    - The subjects must be representative of the target population
    - All scores should be recorded (covering all range of values)
    - We must collect as much as possible of genuine samples and impostors
    - Never assume some parametric form of distribution!

Practical example

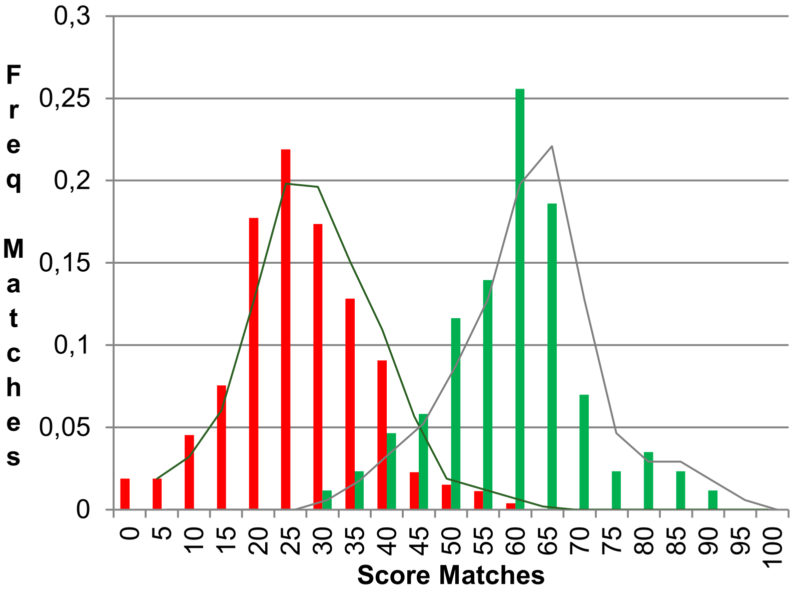
- Example: 10 impostors; 2 legitimates; more then 30 captures of each

Scores	0	5	10	15	20	25	30	35	40	45	50	55	60	65	70	75	80	85	90	95	100	
Attacks	5	5	12	20	47	58	46	34	24	6	4	3	1	0	0	0	0	0	0	0	0	265
Legitimates	0	0	0	0	0	0	1	2	4	5	10	12	22	16	6	2	3	2	1	0	0	86
Attacks	0,02	0,02	0,05	0,08	0,18	0,22	0,17	0,13	0,09	0,02	0,02	0,01	0	0	0	0	0	0	0	0	0	1
Legitimates	0	0	0	0	0	0	0,01	0,02	0,05	0,06	0,12	0,14	0,26	0,19	0,07	0,02	0,03	0,02	0,01	0	0	1
FMR	1	0,98	0,96	0,92	0,84	0,66	0,45	0,27	0,14	0,05	0,03	0,02	0	0	0	0	0	0	0	0	0	
FNMR	0	0	0	0	0	0	0,01	0,03	0,08	0,14	0,26	0,4	0,65	0,84	0,91	0,93	0,97	0,99	1	1	1	
TMR	1	1	1	1	1	1	0,99	0,97	0,92	0,86	0,74	0,6	0,35	0,16	0,09	0,07	0,03	0,01	0	0	0	
User 1	0	0	0	0	0	0	1	2	4	5	7	6	7	4	2	1	1	0	0	0	0	40
User 1	0	0	0	0	0	0	0,03	0,05	0,1	0,13	0,18	0,15	0,18	0,1	0,05	0,03	0,03	0	0	0	0	1
User 2	0	0	0	0	0	0	0	0	0	0	3	6	15	12	4	1	2	2	1	0	0	46
User 2	0	0	0	0	0	0	0	0	0	0	0,07	0,13	0,33	0,26	0,09	0,02	0,04	0,04	0,02	0	0	1



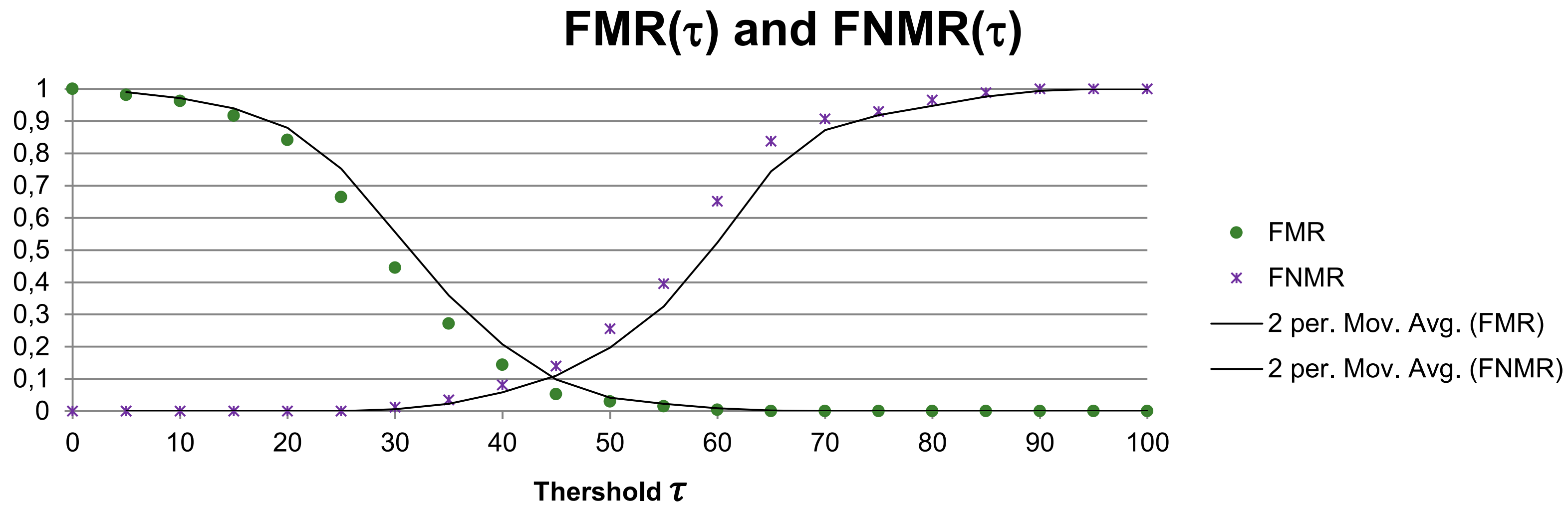
- Performance indicators $\tau = 42$:
  - FM = 14 ⇒ FMR = 0,05
  - FNM = 7 ⇒ FNMR = 0,08
  - TM = 79 ⇒ TMR = 0,92
  - TNM = 251 ⇒ TNMR = 0,95

Practical example – Frequency distribution



- 🟥 Attackers
- 🟩 Legitimates
- – 2 per. Mov. Avg. (attackers)
- – 2 per. Mov. Avg. (legitimates)

Practical example (DET curves)



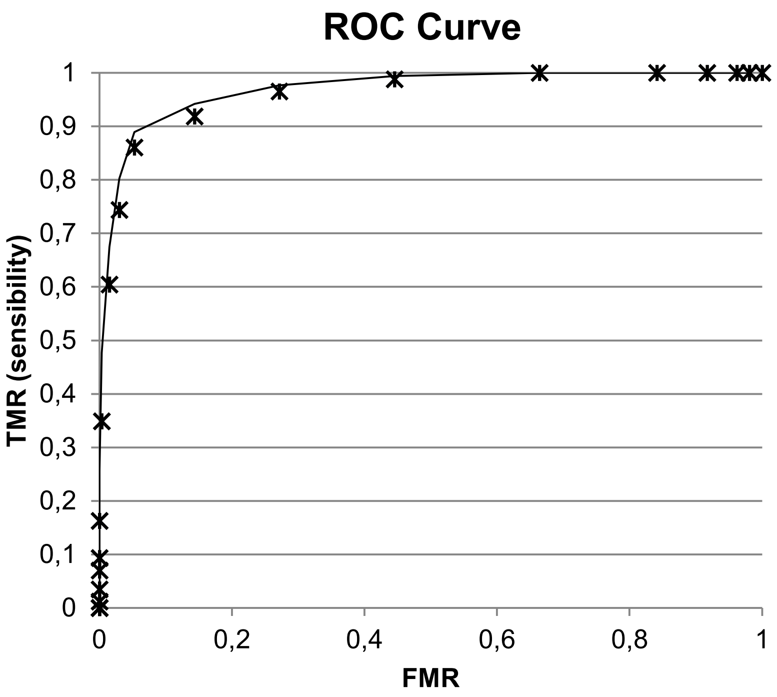
$$\tau_1 = \max_{\tau}\{\tau | FNMR(\tau) \leq FMR(\tau)\},$$

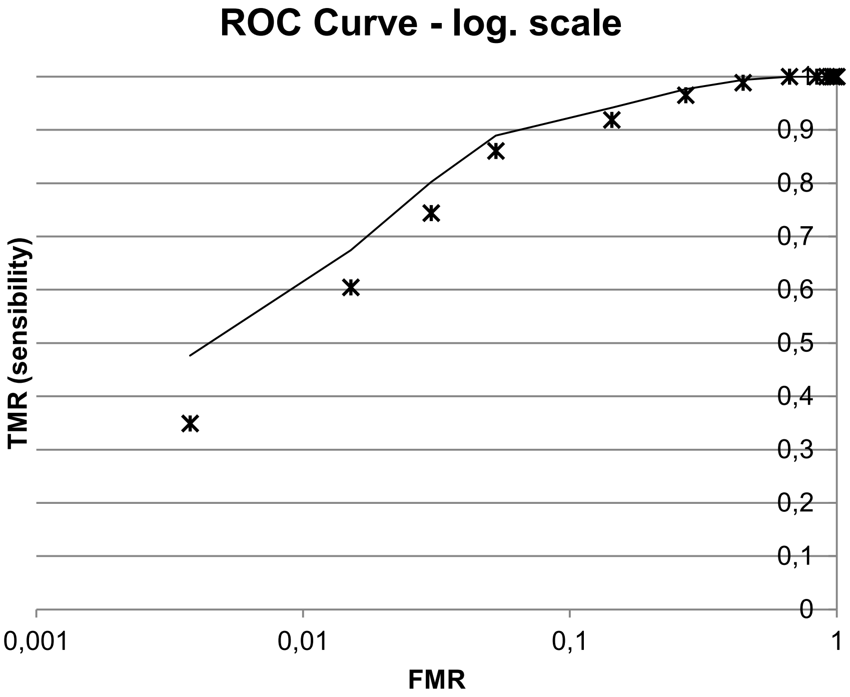
$$\tau_2 = \min_{\tau}\{\tau | FNMR(\tau) \geq FMR(\tau)\},$$

$$[EER_{low}, EER_{high}] = \begin{cases} [FNMR(\tau_1), FMR(\tau_1)] & if\ FNMR(\tau_1) + FMR(\tau_1) \leq \\ & \quad FMR(\tau_2) + FNMR(\tau_2) \\ [FNMR(\tau_2), FMR(\tau_2)] & otherwise \end{cases}$$

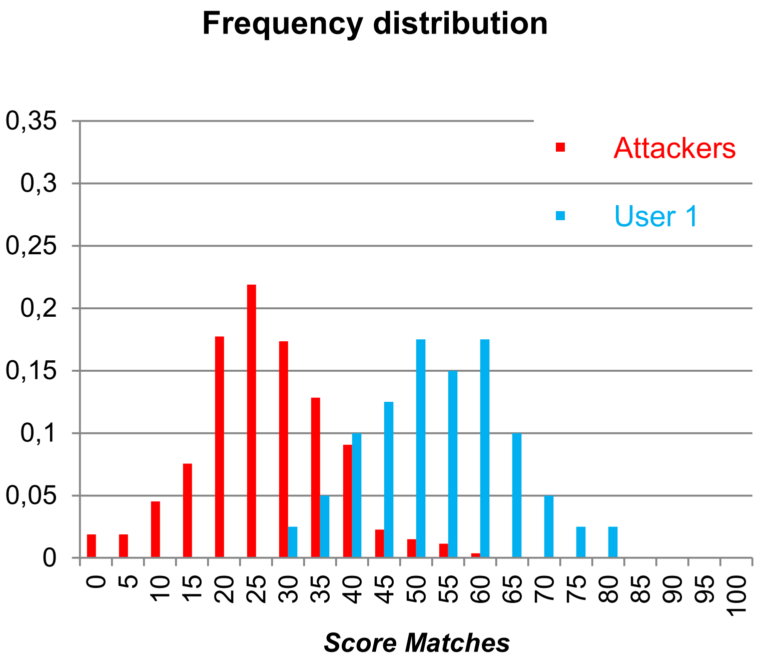
$$e\ EER = \frac{EER_{low} + EER_{high}}{2}$$

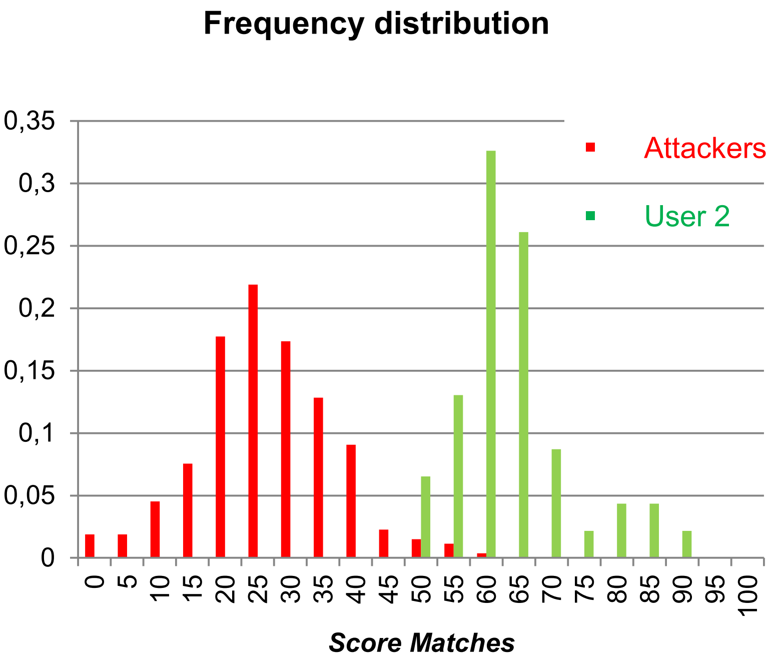
Practical example – ROC curves





Practical example – Individual analysis





Biometrics - precision

- Some indicative values

Typical practical values

Biometry	FTE %	FFR %	FAR %
Fingerprint (FVC [2006])	4	2,2	2,2
Fingerprint (FpVTE [2003])		0,1	1
Face (FRVT [2006])		0,8-1,6	0,1
Iris (ICE [2006])	7	1,1-1,4	0,1
Voice (NIST [2006])	1	5-10	2-5

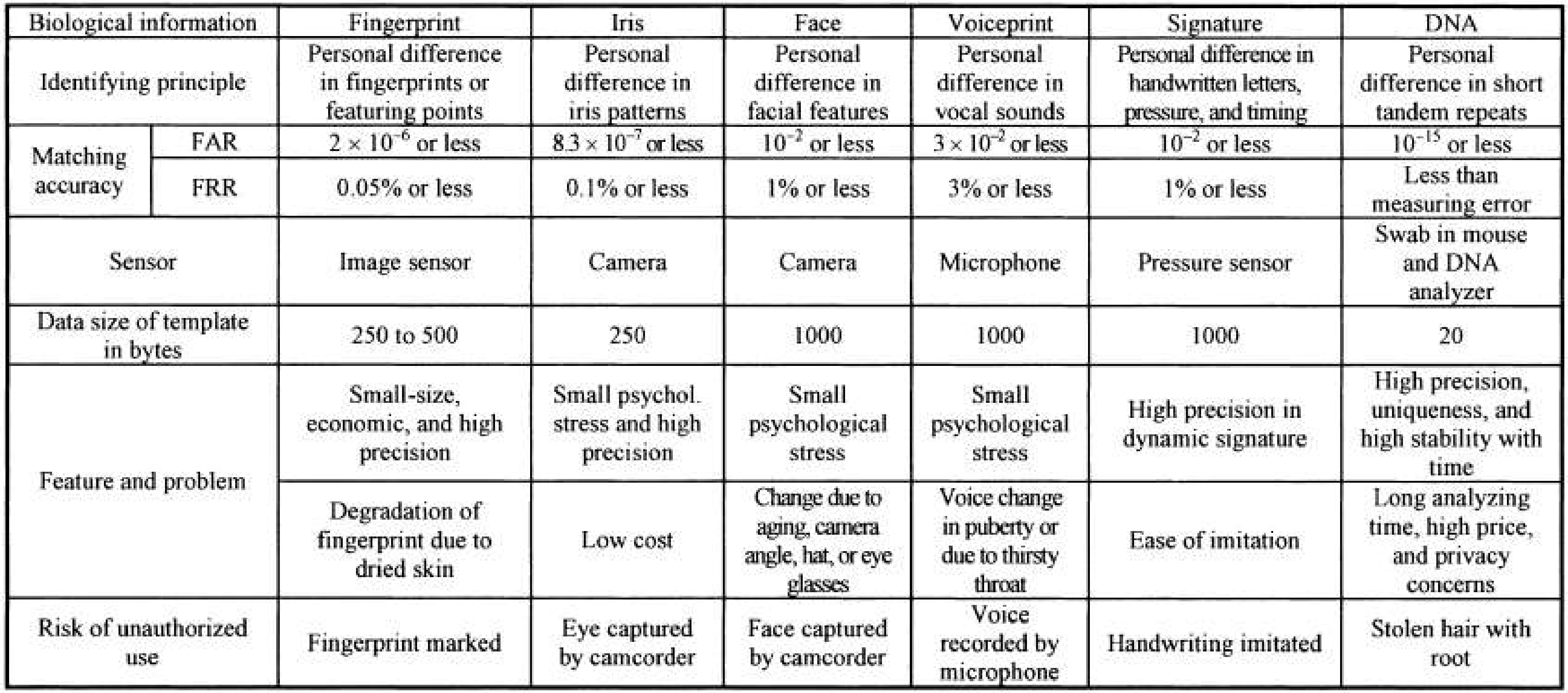
Desirable values

Application	FRR %	FAR %
Authentication	0,1	0,1
Identification (large scale)	10,0	0,0001
Detection	1,0	0,0001

Biometrics - performance

- Other factors relating to performance
  - FTE (Failure To Enroll): number of failures in the registration process
  - FTC (Failure To Capture): number of failures in capturing biometrics
  - Limitations of biological information, inherent to the method
  - Coding limitations
  - Limitations of the invariants (often due to the use of a limited set of test data and learning)

Biometrics in greater detail



Biological information		Fingerprint	Iris	Face	Voiceprint	Signature	DNA
Identifying principle		Personal difference in fingerprints or featuring points	Personal difference in iris patterns	Personal difference in facial features	Personal difference in vocal sounds	Personal difference in handwritten letters, pressure, and timing	Personal difference in short tandem repeats
Matching accuracy	FAR	2×10^{-6} or less	8.3×10^{-7} or less	10^{-2} or less	3×10^{-2} or less	10^{-2} or less	10^{-15} or less
	FRR	0.05% or less	0.1% or less	1% or less	3% or less	1% or less	Less than measuring error
Sensor		Image sensor	Camera	Camera	Microphone	Pressure sensor	Swab in mouse and DNA analyzer
Data size of template in bytes		250 to 500	250	1000	1000	1000	20
Feature and problem		Small-size, economic, and high precision	Small psychol. stress and high precision	Small psychological stress	Small psychological stress	High precision in dynamic signature	High precision, uniqueness, and high stability with time
		Degradation of fingerprint due to dried skin	Low cost	Change due to aging, camera angle, hat, or eye glasses	Voice change in puberty or due to thirsty throat	Ease of imitation	Long analyzing time, high price, and privacy concerns
Risk of unauthorized use		Fingerprint marked	Eye captured by camcorder	Face captured by camcorder	Voice recorded by microphone	Handwriting imitated	Stolen hair with root

Biometrics - scalability

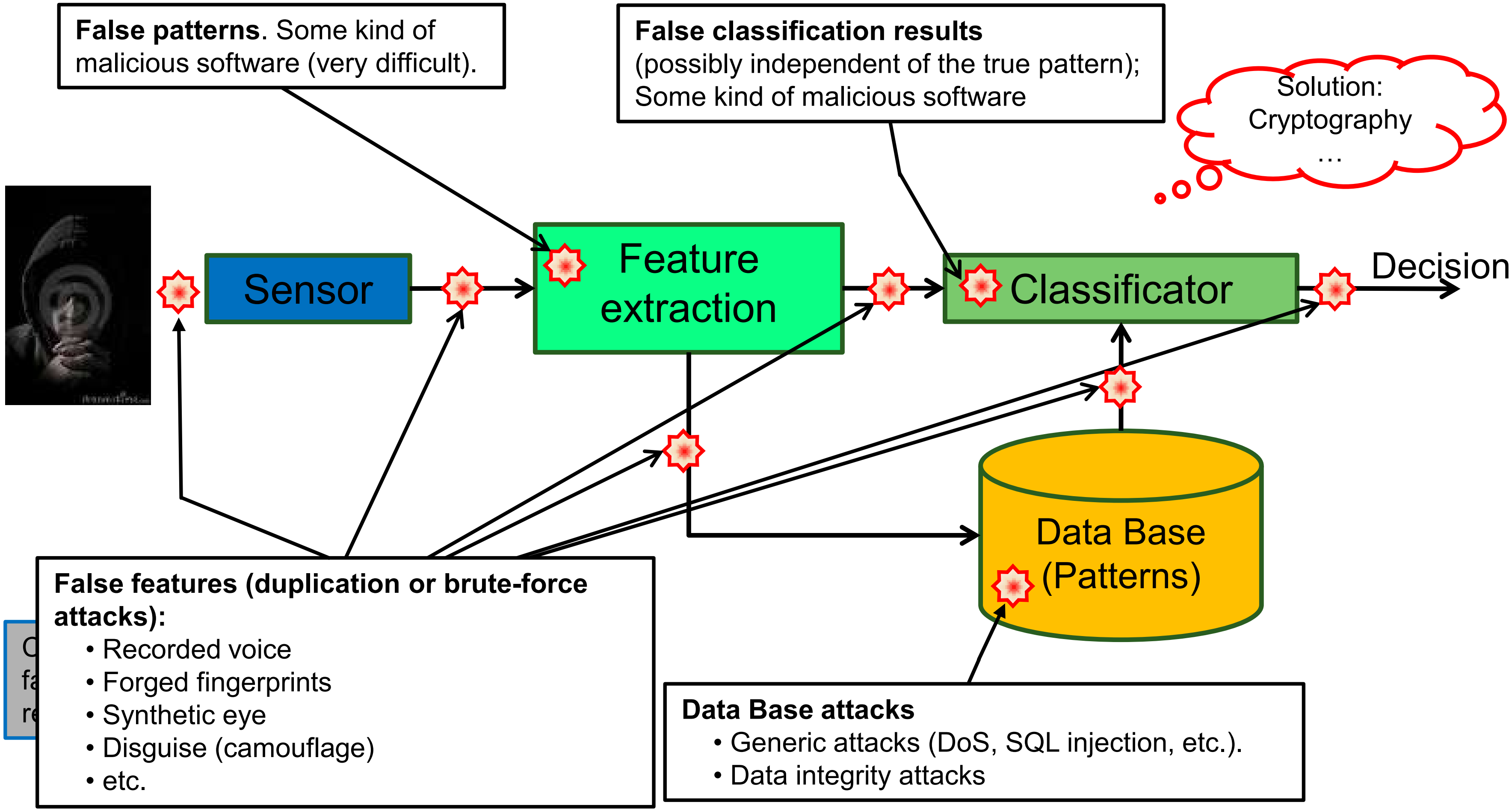
- To what extent the number of individuals enrolled affect system performance?
  - Verification (no problem, since it is an operation **1:1**)
  - Large-scale identification and detection
    - It is not feasible to do **N** operations **1:1**
  - Solutions
    - Adding more computational resources ☹
    - Classification of patterns with exogenous data
    - Verification algorithms more complex and efficient
    - Solutions based on the latter two alternatives tend to have negative impact in performance ☹

Biometrics - Security

- Facts
  - Biometric information is not secret
  - Biometric patterns are not refutable
- Attacks
  - It is (or will be) "possible" to duplicate biometric patterns
  - It is very difficult for the legitimate possessor of a biometric pattern to refute his/her involvement in an attack
  - “Bio-exclusion”
  - Infrastructure Technology Support
- Solutions
  - Ensure “live“ users only
  - **Multi-modal** systems

A consequence of the uniqueness property

Biometrics – Security (technology)



False patterns. Some kind of malicious software (very difficult).

False classification results (possibly independent of the true pattern); Some kind of malicious software

False features (duplication or brute-force attacks):

- Recorded voice
- Forged fingerprints
- Synthetic eye
- Disguise (camouflage)
- etc.

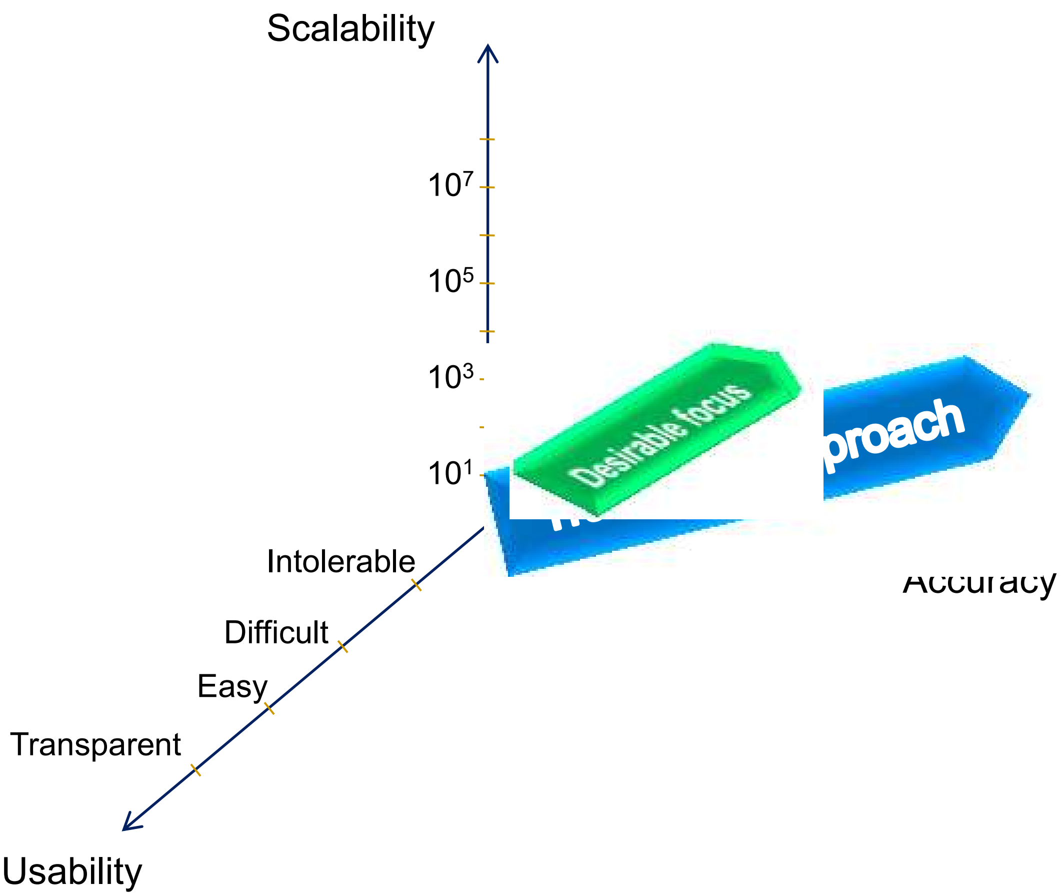
Data Base attacks

- Generic attacks (DoS, SQL injection, etc.).
- Data integrity attacks

Biometrics - privacy

- Biometric data can be used to privacy violation?
- Biometric data can be used for other purposes?
- Biometric data can be used to cross information involving the identity of the individuals?
- Solutions:
  - Biometric Encryption
  - Total transparency
  - Detection systems for "misuse“
  - Multi-modal systems
  - …

Biometrics


Scalability
10^7
10^5
10^3
10^1
Desirable focus
proach
Accuracy
Intolerable
Difficult
Easy
Transparent
Usability

Legal Support

- Law nº 67/98 (personal data; does not specifically mention biometrics)
- Law nº 7/2007 (create the citizen card and governs its deployment and use; does not specifically mention biometrics)
- Working document on biometrics, by the Working Party established by Directive 95/46/EC of the European Parliament: states that biometric data is personal data (general principles)
- The CNPD published:
  - PRINCIPLES ON THE USE OF BIOMETRICS IN THE ACCESS CONTROL AND ASSIDUITY

Conclusions

- Access Control is a key security control
- User authentication is a main issue
- Biometrics: several technologies with high levels of maturity. But ...
  - Scalability is still a problem
  - More research in multi-model biometrics
  - There are no “One Size Fits All” solution
- Usability issues are not solved!
- Computer systems support are often forgotten
- Ability to exploit continuous authentication, enabling the "automatic login"

Bibliografia

- Bishop, M., *Introduction to Computer Security*, Prentice Hall PTR, 2004. (cap 3 a 7)
- Kaufman, C., Perlman, R. and Speciner, M. *Network Security: Private Communication in a Public World*. Prentice Hall PTR, Upper Saddle River, NJ 07458, 2002. (cap. 9)
- Strebe, M., Network Security Foundations: Technology Fundamentals for IT Success: Sybex, 2004. (cap. 3)
- Sandhu, R. S. and Samarati, P., *Access control: Principles and practice*. IEEE Communications Magazine 32(9): 40-49, 1994.
- Sandhu, R. S., E. J. Coyne, et al., Role based access control models, Computer 29(2): 38-47, 1996.
- **Jain, A. K., Ross, A. and Prabhakar, S., *An introduction to biometric recognition*, Circuits and Systems for Video Technology, IEEE Transactions on, 14, 1, 4-20, 2004.**
- Dunstone, T. and Yager, N. *Biometric System and Data Analysis: design, evaluation and data mining*. Springer, 2008.
- Maltoni, D., *et. all, Biometric Systems: Technology, Design and Performance Evaluation*, Springer, 2005
- Jain, A. K., *Biometrics: Personal Identification in Networked Society*, Kluwer Academic Publishers, 1999
- Delac, K. and M. Grgic (2004). A survey of biometric recognition methods. 46th International Symposium Electronics in Marine, ELMAR-2004.
- Yager, N. and Dunstone, T. The Biometric Menagerie. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 32, 2 2010), 220-230
- Magalhães, P. S., *Estudo dos padrões de digitação e sua aplicação na autenticação biométrica* Master Thesis, Departamento de Sistemas de Informação, Universidade do Minho, 2005.
- A. Rashed and H. Santos, Multimodal Biometrics and Multilayered IDM for Secure Authentication, in Global Security, Safety, and Sustainability. vol. 92, S. Tenreiro de Magalhães, et al., Eds., ed: Springer Berlin Heidelberg, 2010, pp. 87-95.

- The Biometric Consortium http://www.biometrics.org/index.htm
- Biometrics Research Homepage at MSU http://biometrics.cse.msu.edu/index.html
- IBG BioPrivacy Initiative http://www.bioprivacy.org/