Security in communications and networks, the issue of (electronic) identity

PA018

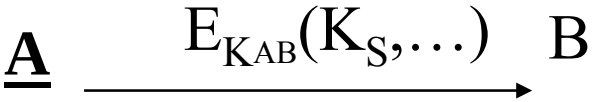
Vašek Matyáš

Major security enablers – critical (infrastructure) applications

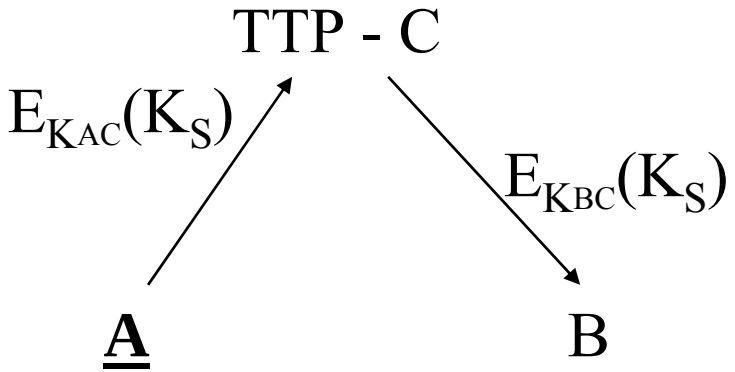
- Kerberos
- Public-key crypto based – certificates, typically X.509 – SSH, SSL/TLS
- Shared-key crypto based – symmetric key ciphers, hash functions

Key distribution (with indirect authentication)

- Direct distribution

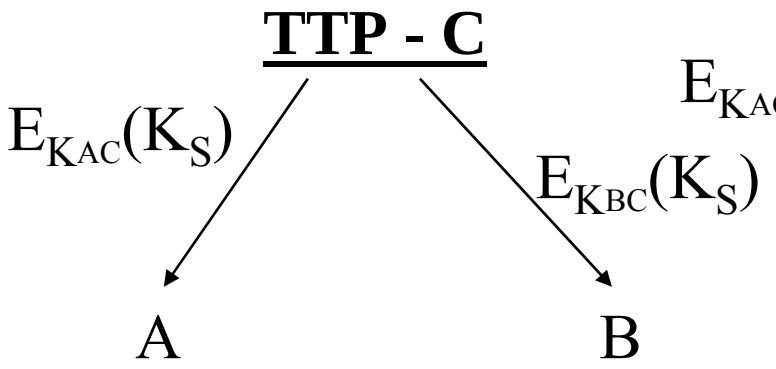


- Key distribution center (also generates the key – following slide)
- Key transport center

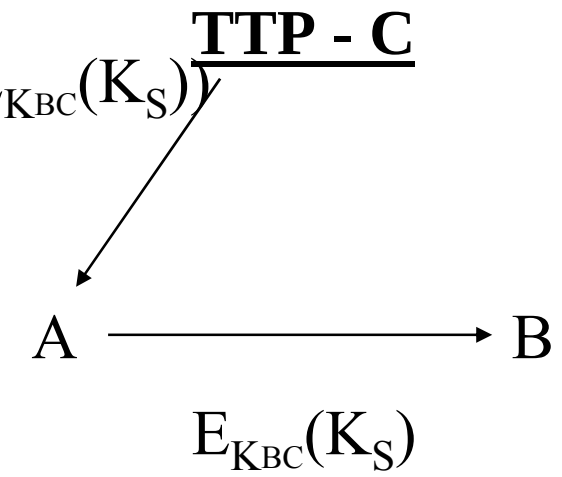


Indirect authentication – key distribution topologies.

TTP-managed



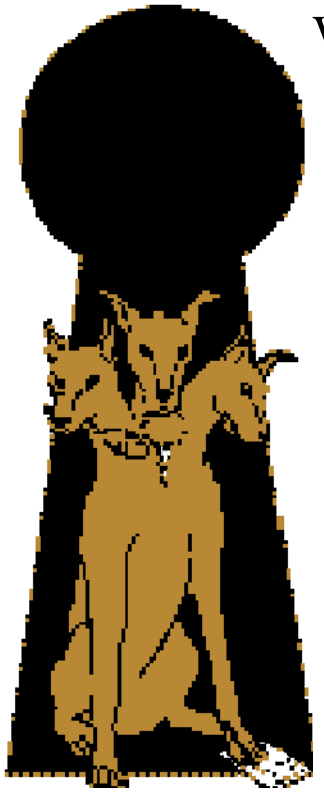
Direct (pull/push)



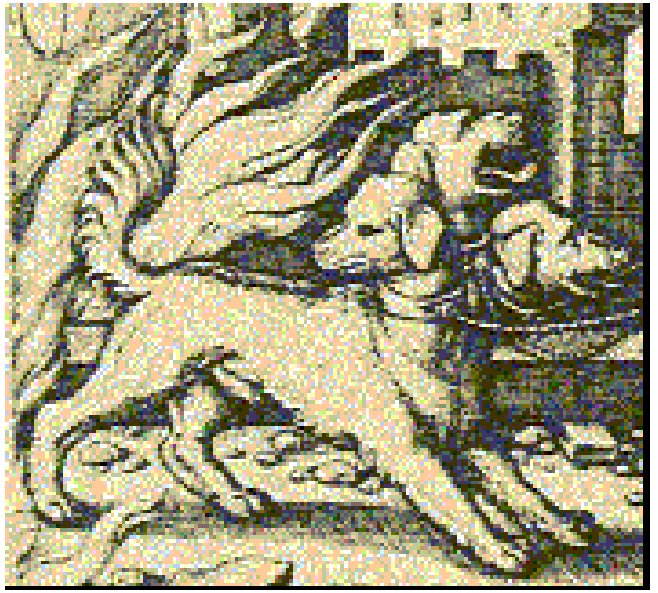
Kerberos

• Greek mythology – guardian to the entrance of Hades (master of the Underground)

• MIT project Athena – MIT's UNIX-based campus-wide academic computing facility



M12.1 Kerberos & Heraklas



Kerberos – threat model

- Users reading messages of other users
- Users replaying messages of other users
- Users altering a workstation network address
- Users impersonating themselves

Kerberos – approach

- Centralised authentication server authenticating both users and machines
- Using symmetric-key techniques, no public-key techniques

Kerberos

- Trusted third-party authentication service
- Key Distribution Center (KDC) grants authentication tokens (“tickets”) to users
  - Trusted, dedicated machine
- Applications can use Kerberos for:
  - Data authentication
  - Data integrity
  - Data confidentiality

Kerberos used for applications

- telnet, rlogin, rcp, FTP, etc.
- Use Kerberos Protocol to exchange authentication information
- Client application uses Ticket-Granting-Ticket to obtain service tickets from KDC
- May use session key to encrypt data checksums (data integrity) or encrypt data (data confidentiality)

Kerberos – important terms

- C = Client
- AS = authentication server
- V = server
- ID_C = identifier of user on C
- ID_V = identifier of V
- P_C = password of user on C
- AD_C = network address of C
- K_V = secret encryption key shared by AS an V
- TS = timestamp
- || = concatenation

Kerberos – simple authentication

- $C \rightarrow AS$: $ID_C \,||\, P_C \,||\, ID_V$
- $AS \rightarrow C$: Ticket
- $C \rightarrow V$: $ID_C \,||\, \text{Ticket}$

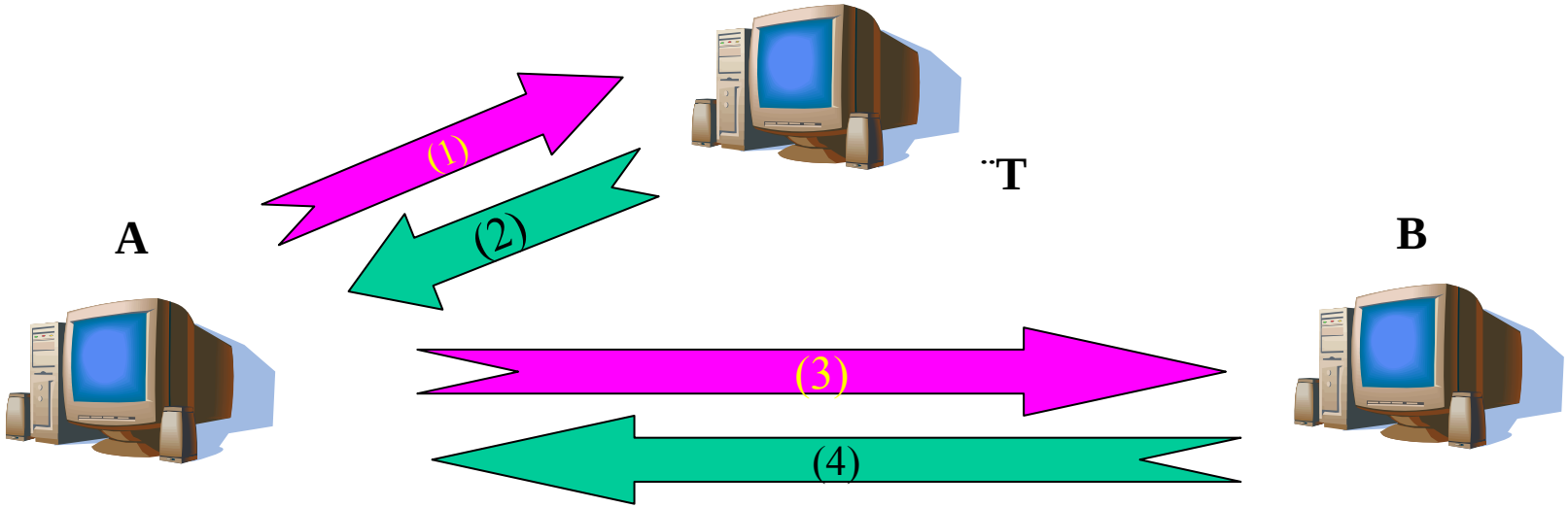
$\text{Ticket} = E_{Kv}(ID_C \,||\, P_C \,||\, ID_V)$

Kerberos Tickets

- Ticket-Granting-Ticket
  - Used to obtain further tickets
  - Requires password or additional authentication from user
  - Lifetime in hours
- Service Tickets
  - Issued to user from KDC
  - User can not decrypt ticket
  - User passes ticket to authenticate to server

Kerberos

- Simplified version of the protocol
  - L – ticket lifetime
  - Def.: $ticket_B = E_{K_{BT}}(k, \text{“A”}, L)$, $auth = E_k(\text{“A”}, T_A)$
  - (1) $A \rightarrow T$: “A”, “B”, n_A
  - (2) $A \leftarrow T$: $ticket_B$, $E_{K_{AT}}(k, n_A, L, \text{“B”})$
  - (3) $A \rightarrow B$: $ticket_B$, $auth$
  - (4) $A \leftarrow B$: $E_k(T_A)$



Kerberos Tickets (Credentials)

- Partly encrypted data structures
  - client ID
  - server ID
  - timestamp
  - session key
  - encrypted part (session key, client info, timestamp)
- Passed the way KDC $\rightarrow$ client $\rightarrow$ server
- Encrypted with the key of intended recipient

Kerberos – time vs. replay issue

- The threat: an opponent steals a ticket and uses it before its expiry time
- Lifetime of the ticket-granting ticket
  - Too short $\Rightarrow$ frequent ticket requests
  - Too long $\Rightarrow$ greater risk of replay attack

Tickets

- Ticket-Granting Ticket – *get once per logon*
- Service-Granting Ticket – *get then once before first use of a service (usually in a given logon session)*

- Authenticated Service Request – *once per (service) session*

Kerberos(v4) Authentication Process

Authentication Service Exhange – To obtain the Ticket-Granting Ticket

1) $C \rightarrow AS$: $ID_C \,||\, ID_{TGS} \,||\, TS_1$

2) $AS \rightarrow C$: $E_{Kc}\,(K_{C,TGS} || \, ID_{TGS} \,||\, TS_2 \,||\, Lifetime_2 \,||\, Ticket_{TGS})$

Ticket-Granting Service Echange – To obtain the Service-Granting Ticket

3) $C \rightarrow TGS$: $ID_V \,||\, Ticket_{TGS} \,||\, Authenticator_C$

4) $TGS \rightarrow C$: $E_{Kc}\,(K_{C,V} || \, ID_V \,||\, TS_4 \,||\, Ticket_V)$

Client/Server Authentication Exhange: To Obtain Service

5) $C \rightarrow V$: $Ticket_V \,||\, Authenticator_C$

6) $V \rightarrow C$: $E_{Kc,v}\,(TS_5 + 1)$

Kerberos today

- **Currently two broadly used versions:**
- 4 - restricted to a single realm (domain)
- 5 - allows inter-realm authentication
- Kerberos v5 is an Internet standard (RFC4120, partly updated by RFCs 4537, 5021)
- MSFT implementation (since Windows 2000)

What is SSL/TLS?

Secure Sockets Layer / Transport Layer Security

- Protocols providing security and reliability
- Protecting communication of two applications
- Running over standard protocols like TCP
- SSL – developed by Netscape, supported also by Microsoft…
- TLS – IETF standard (sometimes called SSL v3.1)
- Transparent for higher-level protocols like HTTP
- Using PKI and X.509 certificates

What security SSL/TLS provide?

Three basic security services:

- **Entity authentication** – the entities are authenticated using server and client certificates.
- **Integrity** – message authentication code (MAC) which ensures the data received is same as the data sent.
- **Confidentiality** – after the initial "handshake", a symmetric key is defined and used to encrypt all subsequent communication (even checked passwords, etc.).

Concepts of SSL/TLS

- Record Protocol
  - –The basic layer of the protocol.
  - –Works over TCP/IP (or other transport protocol).
  - –Allows for encapsulation of different higher level protocols (HTTP, FTP, telnet, etc.) which run unmodified.
- Handshake Protocol
  - –Allows the server and client to authenticate each other.
  - –By default, server authentication is mandatory, client authentication optional.
  - –Authentication through presentation of digital certificates.
    - And verification of the ability to use the related private key!

… more detail

- Establish Session
  - Send random challenge value, accept public key.
  - Verify signed challenge.
  - Deliver session key protected by recipient's public key.
- Communicate Protected Data
  - Encrypt data using agreed cipher and the session key.
  - Produce hash regularly to protect integrity.
  - Data packed into sequenced records.
- (Change Cipher - optional)
- Finish Session
- *http://tools.ietf.org/html/rfc5246 (partial update by RFC5746 in Feb 2010)*

Intrusion Detection Systems

- Intrusion – activity aimed at disrupting or circumventing a service within an organization's system
  - Not Intrusion Prevention(?) Systems ☺
- Also penetration, breach (, attack)
- Technical methods
  - Not social engineering

IDS Principles

- Anomaly detection
  - Unusual pattern (as compared to typical user/system behavior).
  - False positives!
- Misuse detection
  - Pattern of intrusion(-like) behavior
  - False negatives!

Combine these two approaches!

IDS Topologies

Network-based

- Checking network traffic
- Use raw network packets.
- Typically a network adapter running in promiscuous - monitoring and analyzing all traffic.
- Responses like admin notification, connection termination, session recording (for forensic analysis), other detailed evidence collection.

Host-based

- Checking machines (log files, etc.).
- Started in 80s – log file review.
- Typically monitor system, event, and security logs on WinNT and syslog on Unix.
- Also critical file checksum control, response time, port activities.
- Responses analogous…

Combine these two approaches!

Secure SHell

- SSH
- `http://www.ssh.com/`
- Non-commercial downloads

- WinSCP
- `http://winscp.sourceforge.net/eng/`
- WinSCP

Email Security

- Postcard-like service

- (X.400)
- PGP (Pretty Good Privacy)
- S/MIME (Secure Multipurpose Internet Mail Extension)

S/MIME messages

- Combinations of two separately defined formats
  - (1) MIME entities
  - (2) Cryptographic Message Syntax (CMS) objects
- S/MIME entity formats
  - **one** for **enveloped** (i.e., encrypted) – provides confidentiality and key distribution services
  - **two** for **signed** – each provides integrity and data origin authentication services
  - **nested combinations** of signed and encrypted formats
  - may nest in any order to any “reasonable” depth
  - multiple nesting is used to construct S/MIME Enhanced Security Services

Firewalls

- Protect against attacks from the outside (across the firewall)
- Attacks against internal data
- Denial-of-service attacks
- Communication options:
  1. Allow
  2. Deny
  3. Translate (Proxy)

Basic options – firewalls

TCP/UDP	Allow/Deny	Packet filtering *(routers)*
TCP	A/D/Translate	Circuit-gateway *(trust inside)*
HTTP, FTP…	A/D/T	Applic.-gateway

Closely related topics – to be discussed later.

- Firewalls and network security
  - Guest lecture next week – Josef Pojsl, Technical Director, Trusted Network Solutions

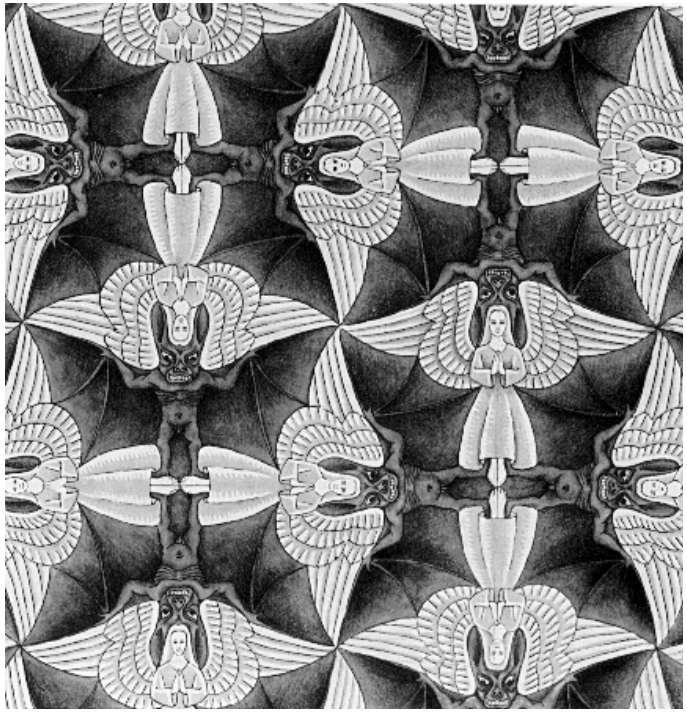
Identity

in computer and communication systems

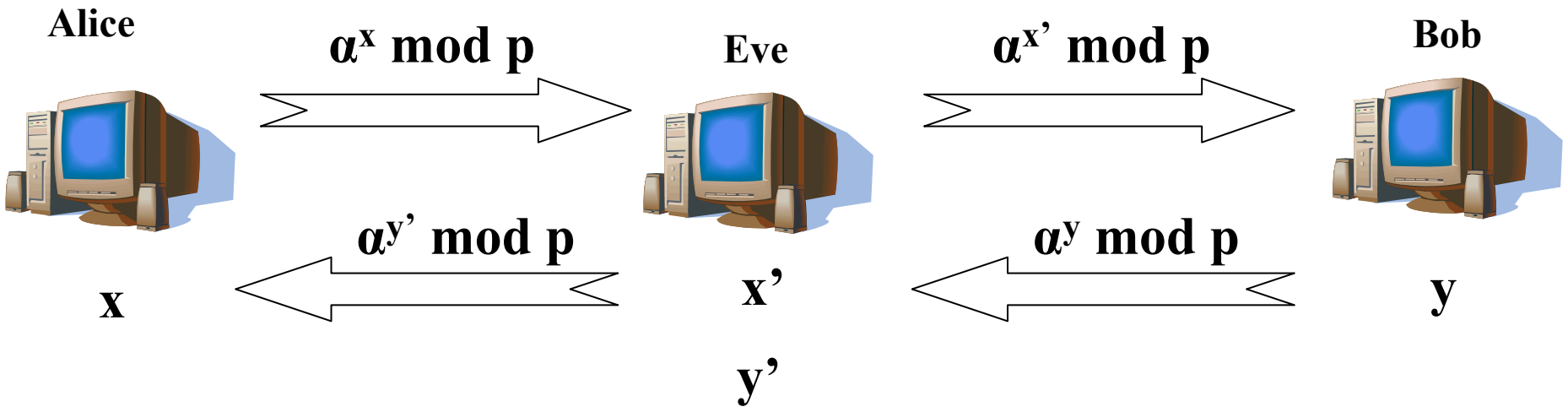
M.C. Escher,
also later in the talk

Agenda

- Learning from mistakes
- Authentication in computer systems
- Identity – personal, in computer systems
- Information privacy
- Case 1 – Passwords
- Case 2 – User control
- Conclusions
- Reading



Diffie-Hellman(-Merkle) protocol, man-in-the-middle attack



Alice believes that she communicates with Bob, and vice versa; Eve reads and possibly can modify passing messages.

To avoid this attack – ensure that they are indeed using each other's public keys.

Needham-Schroeder public-key protocol, man-in-the-middle attack

1. $A \rightarrow E : P_E(N_A, A)$
2. $E \rightarrow B : P_B(N_A, A)$
3. $E \leftarrow B : P_A(N_A, N_B)$
4. $A \leftarrow E : P_A(N_A, N_B)$
5. $A \rightarrow E : P_E(N_B)$
6. $E \rightarrow B : P_B(N_B)$

B believes that

- He communicates with A
- N_A and N_B are known only to A and B.

To avoid this attack, B has to be more explicit in step 3(2), i.e. $P_A(N_A, N_B, B)$. Attack due to G. Lowe, 1995.

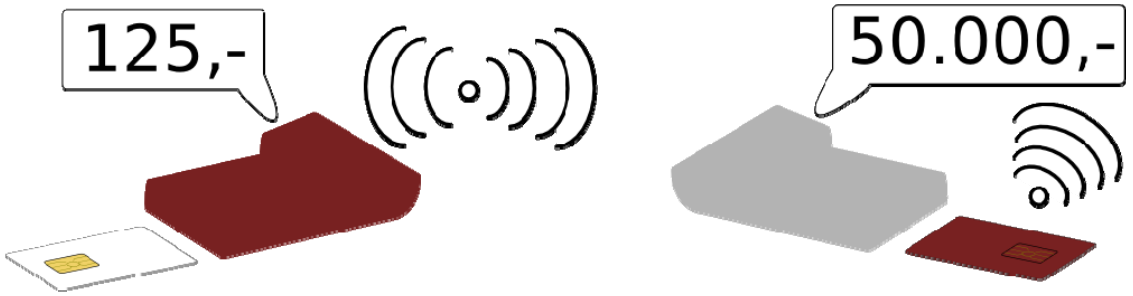
“Attacks” on SSL

- Man-in-the-middle is an evergreen
  - Most recent 2009, due to Moxie Marlinspike
  - Build often on poor check of public-key certificates by users
  - …or problems with inconsistent public-key certificate check by browsers or servers
  - …or favorite icon display in the URL bar 
  - …or abusing layers of indirection (HTTP to HTTPS)
- Public-key certificates overloaded – attribute certificates
- Issues beyond technology – adequate precautions, from both a legal and a personal view

Recent attacks on Chip & PIN

Problem of untrustworthy terminal – authentication failure

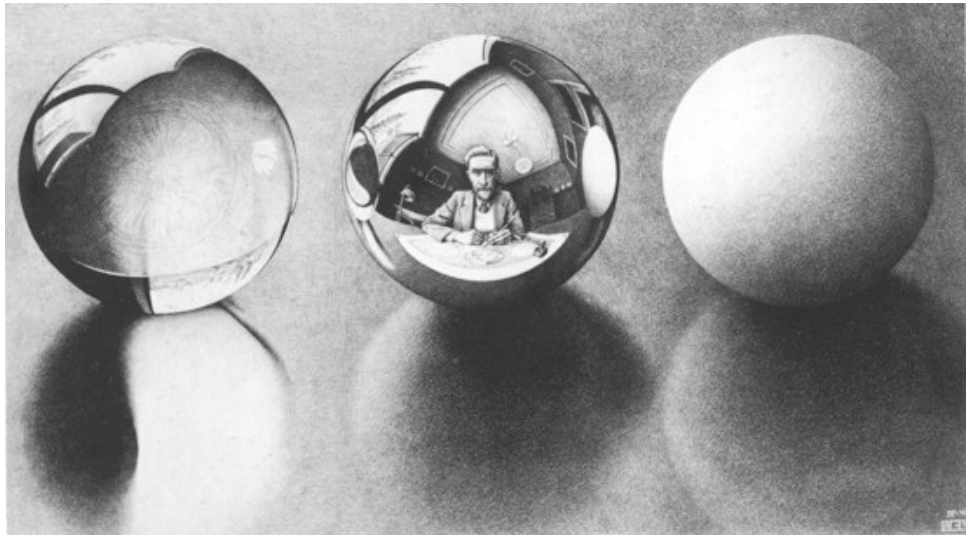
- either with unauthorized wireless broadcast



- or with unauthorized device between the card and reader
  - device → PINpad : card authentication check OK
  - card → device → PINpad : cryptogram indicating PIN check failure
  - PINpad → bank : card auth. check OK, cryptogram with PIN check failure
  - bank → PINpad : sale is OK (signature authorization assumed)

Agenda

- Learning from mistakes
- Authentication in computer systems
- Identity – personal, in computer systems
- Information privacy
- Case 1 – Passwords
- Case 2 – User control
- Conclusions
- Reading



We grow…

- Every year we add
  - Nearly 200 million new "standard" computers
  - Over 50 million cars
    - Average car has got about 50 CPUs built in
  - About 1 billion new mobiles
  - Over 5 billion chip cards (almost 90% with CPU)
  - Add PDAs, e-passports and other RFIDs, sensor nodes, trains, planes, home appliances, …
- Mobile phone subscribers – 4.5 billion (2009)
  - 1 billion in 2002
  - GSM networks operate in more countries and other territories than the UN recognizes (192 cf. 219)

Entity authentication

- Differs from message/data authentication
  - Timeliness guarantee for entity authentication
    - Claim/verification of identity in real-time
    - Importance of time-variant parameters
  - Transferred data is of little value afterwards
- Unilateral / mutual
- Secret-based authentication
  - Weak
  - Challenge-response
  - Zero-knowledge

Knowledge of secret key ⇒ identity

- For shared-key crypto based on
  - trust in the party the key is shared with
  - *Authentication ~ Ability to en-/de-crypt or MAC*
- For public-key crypto based on
  - trust in the party possessing the private key and
  - trust in link between the public key and other data
  - *Authentication ~ Ability to sign or decrypt messages*

Entity authentication protocol

1. At least one of the honest parties is able to successfully authenticate itself
2. The verifier cannot reuse the authentication exchange to impersonate the claimant to 3rd party
3. Negligible probability of an attacker to play the role of the claimant
4. Point 3 is true if is a (polynomially) large number of past exchanges has been observed by the attacker

Biometrics – Identification vs. Authentication

Determination of a person's identity. (1:N)

"Positive authentication"

Hard to achieve

- Small user groups.
- Low accuracy.
- Exception: iris scan.

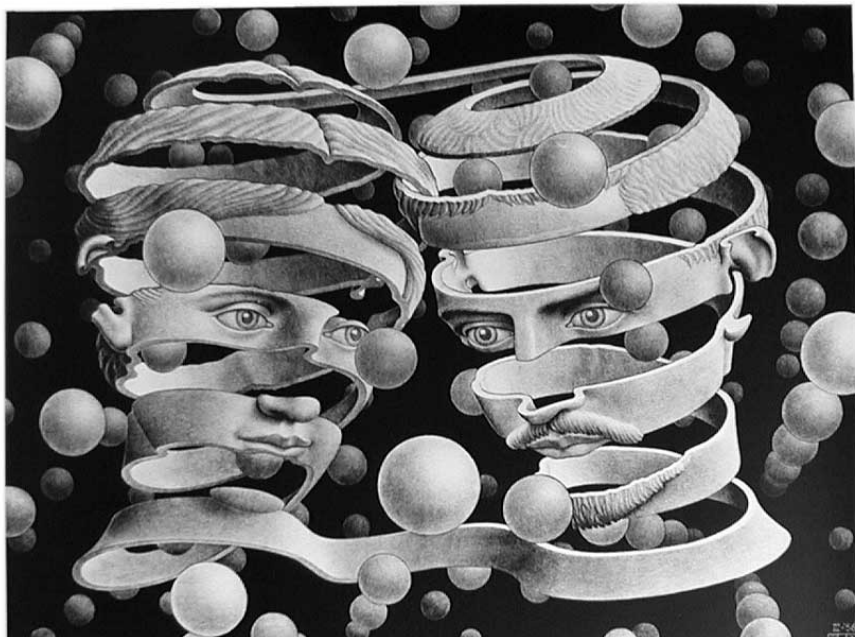
Verification of a person's identity *claim*. (1:1)

Easier than identification.

User group size – accuracy!

Agenda

- Learning from mistakes
- Authentication in computer systems
- Identity – personal, in computer systems
- Information privacy
- Case 1 – Passwords
- Case 2 – User control
- Conclusions
- Reading



Identity

- Multidisciplinary challenge – Philosophy, Law, Technologies, Social Sciences, Mathematics, Biology, Informatics…
- Major evolutions
  - Social changes (war, taxes and state, travel)
  - Law (virtual persons came well ahead of IT – *Nondum Conceptus*, *Nascirtus*)
  - Technology (data processing, Internet, ambient intelligence)
- *Idem* – sameness asserted in difference to others – similarity and continuity
- *Ipse* – selfhood – human subject, I/me

Personal identity

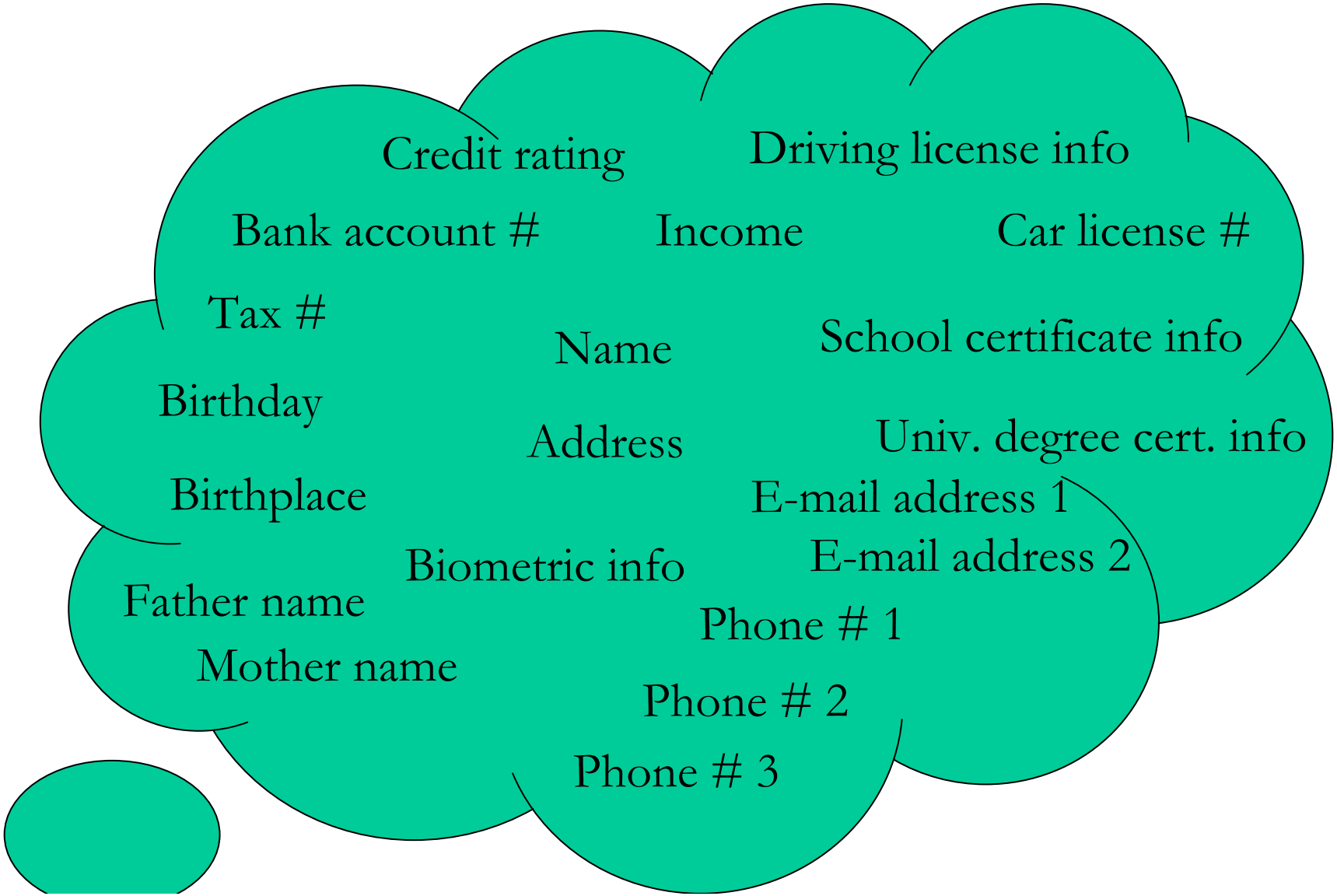
  - Biological
  - Psychological
  - Social
- Criminology assumes that identity of a person does not change with time
- Identification
  - Internal
  - External
- Need for better identification – surnames … ID #s
  - Shanghai with 8 mil. people using 408 surnames, all China registering 3,100 surnames and Chinese Top 5 (Zhang, Wang, Li, Zhao, Chen) used by 350 million people

Wikipedia – Identity – Computer Sci.

- Identity – *object-oriented programming* – describes the property of objects that distinguishes them from other objects
- *Identity column* – database field that uniquely identifies every row in the table and is made up of values generated by the database
- *Federated identity* – assembled identity of a person's user information, stored across multiple distinct identity management systems

- *Digital identity* – representation of a set of claims made by one digital subject about itself or another digital subject
- *Identity management* – administrative area that deals with identifying individuals in a system and controlling access to resources by placing restrictions on them
- *Online identity* – social identity that an internet user establishes in online communities and websites

Multiple facets of identity



Categorization of attributes

- *Domain* – work, education, health, government
- *Functional* – identification, location, social group, biological, psychological-personal
- *Temporal*
  - *Permanent-given* – sex, eye colour, parents, DoB,…
  - *Permanent-acquired* – qualification, behavioural,…
  - *Persistent-situations* – address, marital status,…
  - *Transient* – location, haircut, clothing,…

Problems of Personal Identity (Stanford Encyclopedia of Philosophy)

- Who am I?
- Personhood - What is it to be a person?
- What am I?
- How could I have been?

- Persistence
- Evidence
- Population
- What matters in identity?

Personal identity – critical issues for IT

- Imperfection of the representation of the external view
  - we "reduce" a person to attributes and
  - often then the attributes farther to their digital representation
- Control of the attributes/information
  - Some by the person/subject
  - Some by institutions (government, insurance…)

Lessons from distributed systems

- Pure names (IDs)
  - Of little use in distributed systems
  - One must know where to look them up
  - Directory services become critical
- Centralized systems had some idea of the name set size
- Distributed systems
  - Good design assumes the opposite
  - Assume indefinite # of machines, each with a lookup/directory service of indefinite size
- Prominent feature – measures to avoid (=resolve) confusion by accidental non-uniqueness of naming
- Uniqueness control – hierarchy (divide and conquer)
  - Such hierarchy must reflect reality (mgmt. / communication)

Identity and attributes – issues/crime

- Identity collision – accidental wrong link
- Identity change – intentional wrong link
  - Identity delegation – with consent
  - Identity takeover – without consent
  - Identity exchange
  - Identity creation
- Identity obstruction – link is deleted
- Identity restoration – link is restored

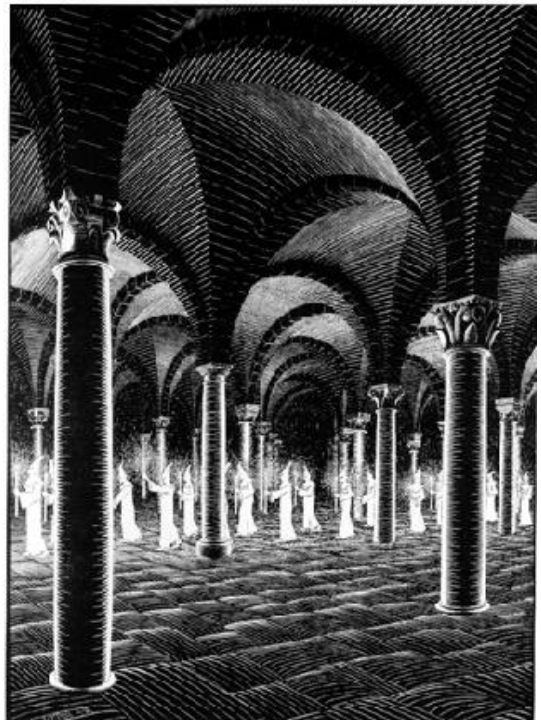
- Identity “theft” = takeover, “fraud” = fraud

Personal data (European legal view)

- Any data concerning identified or identifiable data subject
- Data subject is identified or identifiable if his/her identity can be directly or indirectly determined from one or more personal data (items)
- This holds true only if the effort to determine identity does not consume overly high time, effort or material resources

Agenda

- Learning from mistakes
- Authentication in computer systems
- Identity – personal, in computer systems
- Information privacy
- Case 1 – Passwords
- Case 2 – User control
- Conclusions
- Reading



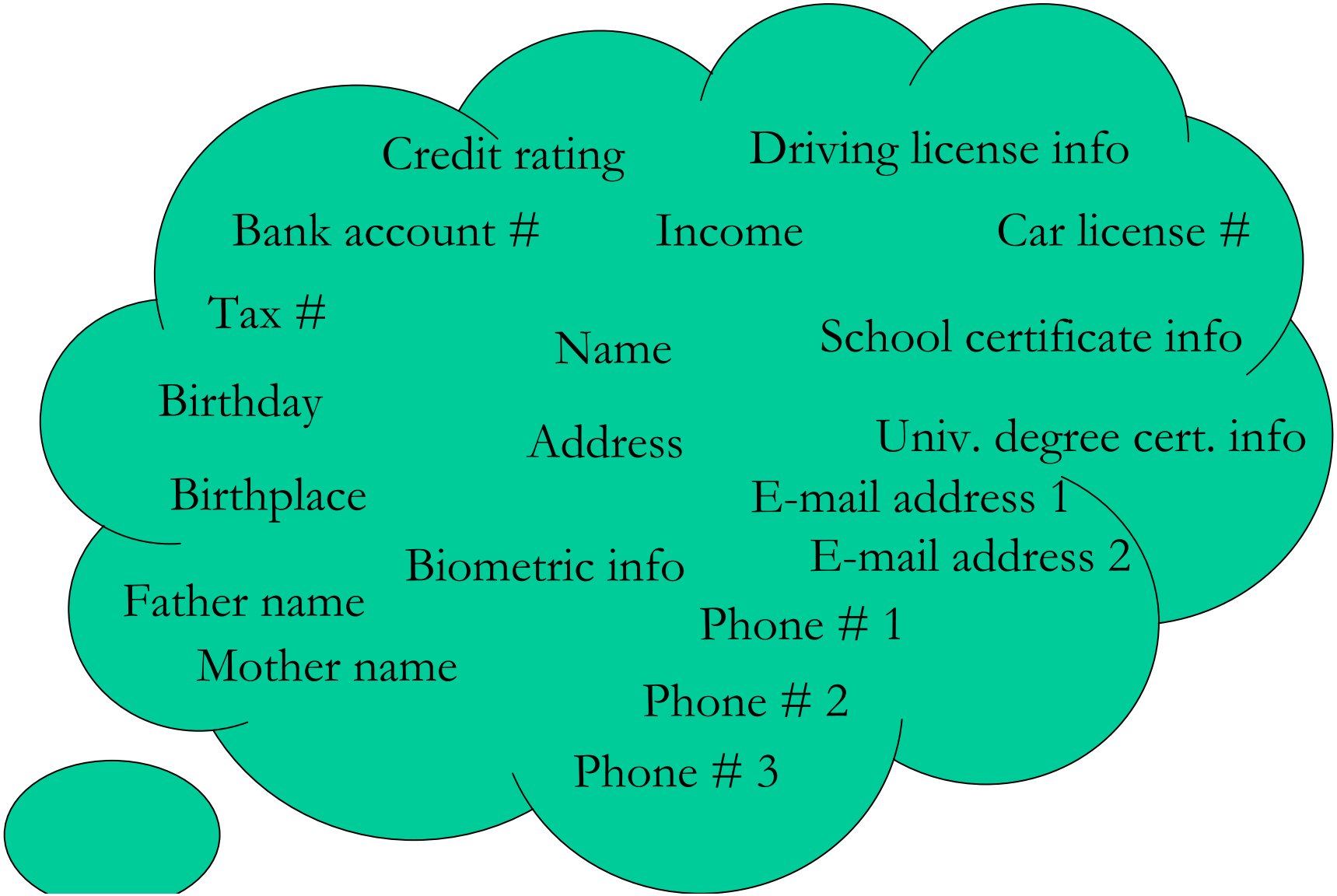
Pfitzmann, Hansen et al. approach

- *A terminology for talking about privacy by data minimization: Anonymity, Unlinkability, Undetectability, Unobservability, Pseudonymity, and Identity Management*
- Gradually (over 9 years) amalgamated document
- Usual message system (sender, network, message, recipient) setting
- Attacker's point-of-view
  - Also with attacker's active participation
  - Yet disregarding the message content
- See Common Criteria (CC) for some comparisons

Identifiability, identity

- Possibility that an attacker can sufficiently identify the subject within a set of subjects, the *identifiability set*
- *Identity* is any subset of attributes of an individual which uniquely characterizes this individual within any set of individuals
  - Usually there is no such thing as "the identity", but several of them
- *Partial identity* – a subset for specific role or context or community

Identity and partial identities



Digital identity

- Attribution of values to an individual person, with immediate operational access to the values by technical means
- *Identity management* – managing partial identities (pseudonyms) of an individual person, i.e., administration of identity attributes

Anonymity, anonymity set

- Anonymity – the state of being not identifiable within a set of subjects, the *anonymity set*
  - Not identifiable = not uniquely characterized within
- *Anonymity set – subset of all subjects who might have undertaken a certain action* (e.g., sent a message)
- Larger anonymity set => stronger anonymity

Pseudonymity

1. Being pseudonymous is the state of using a pseudonym as ID
2. Pseudonymity is the use of pseudonyms as IDs

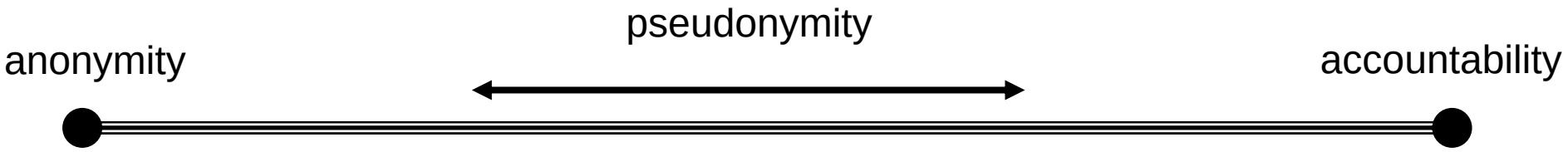
- Greek “pseudonumon” = “falsely named” (pseudo: false; onuma: name)
  - Also a suggestion to consider this being a mapping from “real name” into another name
  - ... tricky(!) – is it the “another name” or the mapping?
- Pseudonymization – de-identification (label change) of data for data protection

Digital pseudonym (accountability)

- A bit string which, to be meaningful in a certain context, is
  - unique as ID (at least with very high probability)
  - suitable to be used – w.r.t. a particular community (size) – to authenticate the holder and his/her action(s), e.g., message(s) sent

- Using digital pseudonyms, accountability can be realized with pseudonyms

Pseudonymity & linkability



- *public pseudonym* (link always publicly known)
- *initially non-public pseudonym* (link initially only known to certain parties)
- *initially unlinked pseudonym* – only holder knows

Reputation & resolving problems:

  - third parties (identity brokers) have a way to reveal the civil identity of the holder in order to provide means for investigation or prosecution

Linking evidence through pseudonyms

- Evidence (context information):
  - Too little – limited trust level being achieved
  - Too much – potential breach of privacy
- Users use a number of pseudonyms (identities/aliases)
  - And reveal links between them as it suits (for a gain of some kind)
- Parties involved can combine their data and so breach privacy of their clients

PATS proposal

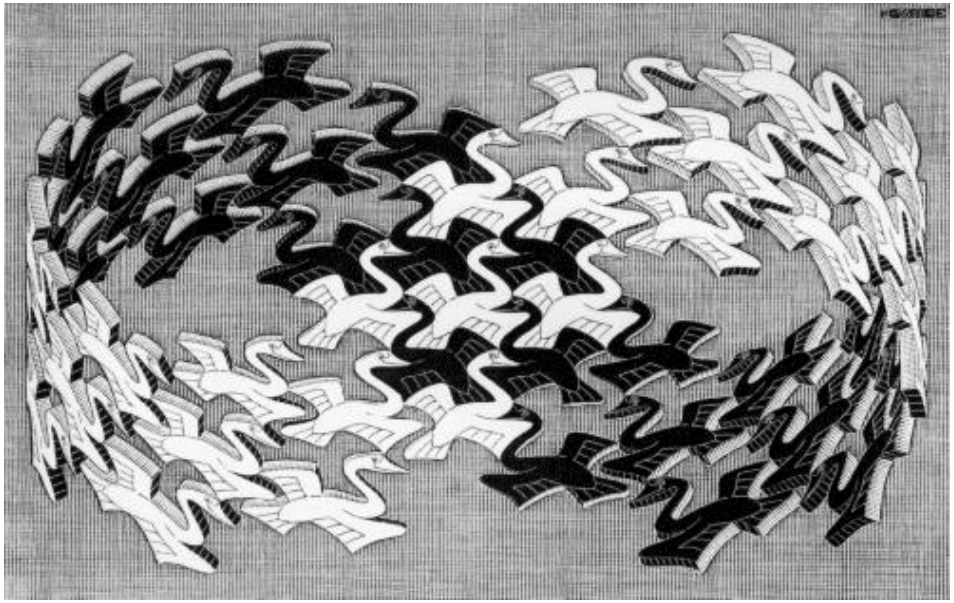
- Goal: find the best link between two vertices
  - e.g., event anonymity – user and event IDs
- The graph represents attacker's knowledge (context info) at a given time
- Introduce all context info as vertices
- Edges – probability weights of vertices' connection/relation (linkability)
- Normalization of the graph
  - User IDs don't link directly (only through other context info)
  - Same for service IDs
  - Edge weights (introducing domains of vertices)

Use of pseudonyms

- More flexibility for both system designer and user than with anonymity
- Can lead to different pseudonyms implied by different sets of evidence
- Yet issue of mutual linking for distinct pseudonyms
  - With temporary links possibly desirable in case of need to achieve higher trust/reputation level. (Temporary in the sense of user privacy protection interest.)
- System parameter: How hard it is to create a new pseudonym with good (enough) reputation?

Agenda

- Learning from mistakes
- Authentication in computer systems
- Identity – personal, in computer systems
- Information privacy
- Case 1 – Passwords
- Case 2 – User control
- Conclusions
- Reading



Passwords

Human memory vs. security
(short easy-to-guess string vs. long complicated string)

- Dictionary attacks
  - Today combinations of up to 8 characters
  - Common words and user-related values, permutations, substitutions, etc.

  - Usual success rate 20-40%

Password quality checker

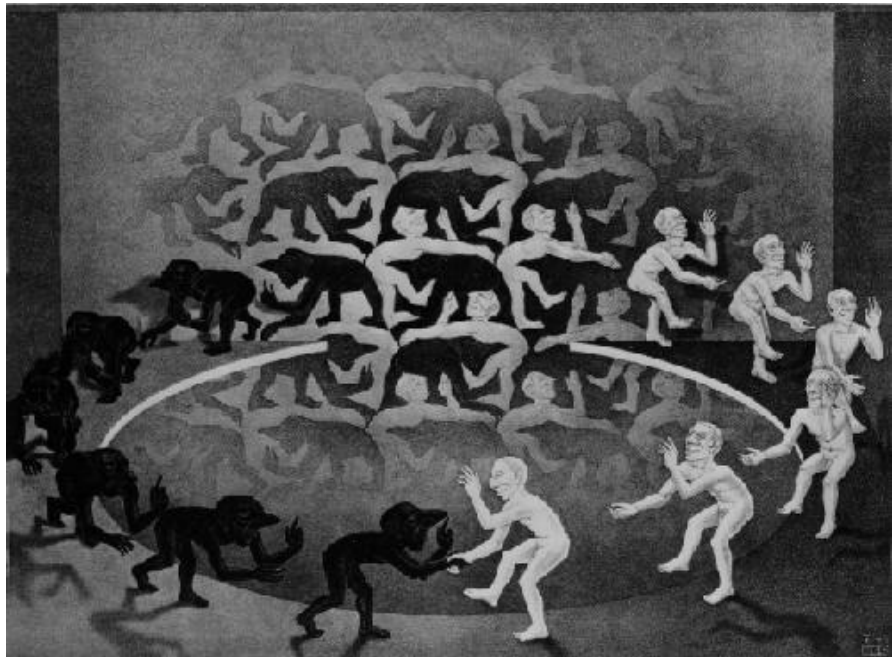
- Critical question: *How good is my password?*
- PGP – meter with instant response (not suitable – for passphrases)
- Securecode.net – online strength meter – after the password is entered
- Lotus Notes (password policy setting, then Y/N responds w.r.t. the setting)

Our design

- Users see *how good* (in 0-100%) their password is – unlike typical Y/N response
  - Password length (set starting value – 0 for 1,… 100 for 8) – then multiply
  - Alphanum. types (0.1 for 1, 1.0 for 4)
  - Distinct chars (0.01 for 1,… 1.0 for 5)
  - Dictionary checks (with modifications, substitutions, etc.) – 0.1 for each "hit"
- Improvement (20-person sample) from 17% to 75% after education & illustration

Agenda

- Learning from mistakes
- Authentication in computer systems
- Identity – personal, in computer systems
- Information privacy
- Case 1 – Passwords
- Case 2 – User control
- Conclusions
- Reading



Payment device interface

- Under customer control
- Enables verification of the transaction independently of the PINpad

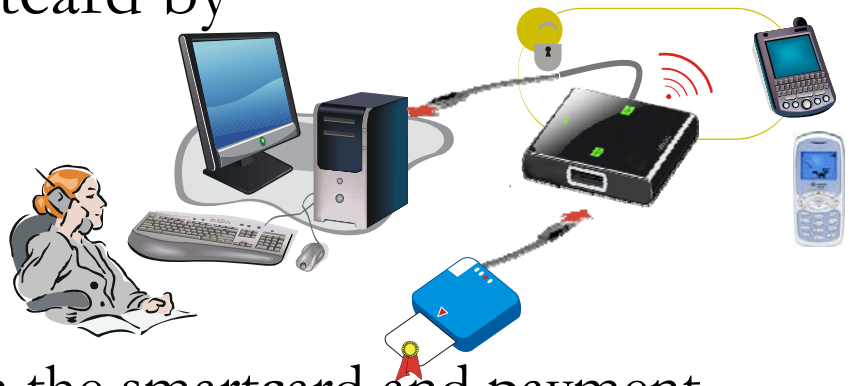
Project “supertoken”

- Requirements:
  - Standard protection of authentication data
  - Secure use in insecure devices (PC, PINpad)

- Architecture:
  - two logical rings of protection:
    - Crypto (smartcard) chip and data storage
    - Independent access control (to the chip)
  - Access control
    - Direct user interaction
    - Independent I/O unit integrated within the token

New token – “check&sign”

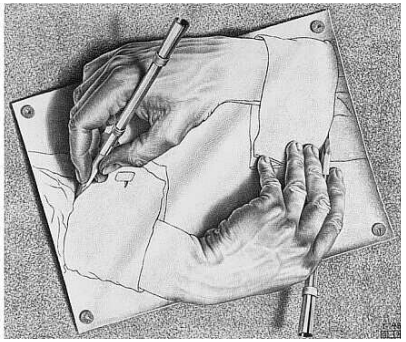
- Supertoken extends the smartcard by
  - USB token
  - External display (PDA, phone)

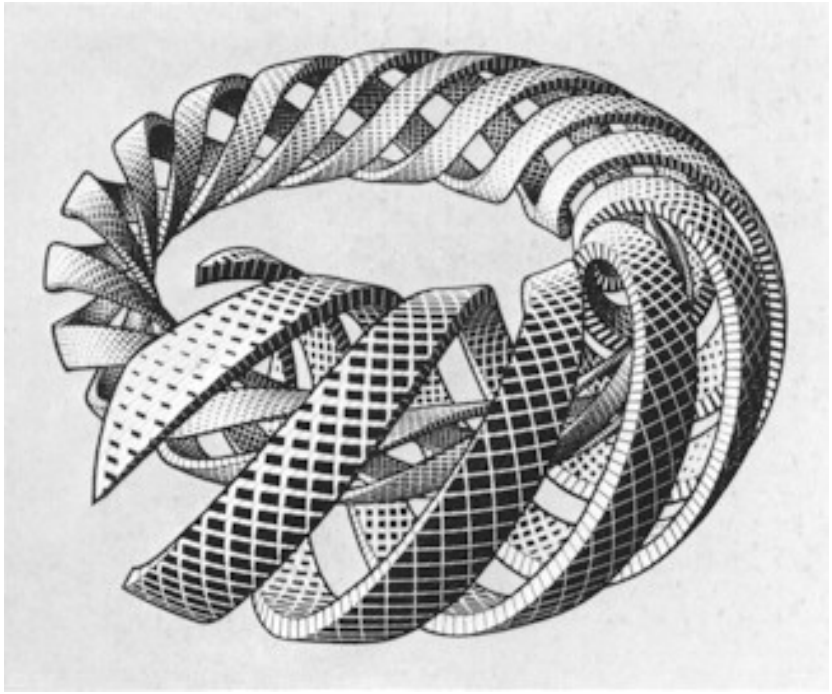


- USB token
  - Controls the data flow between the smartcard and payment device (computer, PINpad)
  - Displays critical transaction info (amount, account #, etc.)
- Customer
  - Checks and confirms the validity of displayed data before the signature (cryptogram) is sent to the device
- Alternative – keyboard integrated “type&sign” version

Conclusions – minimize risks

0. For authentication protocol designers
   - Be as explicit as possible (R Needham) & bind exchanges well
   - Have fallback procedures & resources
1. For system designers
   - Assume that you don't know the # of users
   - Be ready for accidental non-uniqueness of naming
2. Privacy is not only about data confidentiality, but namely about links between data items
3. Users should have (just) enough and reliable information when authenticating & making decisions





Thank you for your attention!

…and I should not forget the reading on the next slide…

Reading for the 3rd and 4th weeks

- A Pfitzmann, M Hansen et al., *A terminology for talking about privacy by data minimization: Anonymity, Unlinkability, Undetectability, Unobservability, Pseudonymity, and Identity Management.*
  - Online at TU Dresden:
  - http://dud.inf.tu-dresden.de/Anon_Terminology.shtml